

---

P L Krapivsky · Kirone Mallick · Tridib Sadhu

# Tagged particle in single-file diffusion

**Abstract** Single-file diffusion is a one-dimensional interacting infinite-particle system in which the order of particles never changes. An intriguing feature of single-file diffusion is that the mean-square displacement of a tagged particle exhibits an anomalously slow sub-diffusive growth. We study the full statistics of the displacement using a macroscopic fluctuation theory. For the simplest single-file system of impenetrable Brownian particles we compute the large deviation function and provide an independent verification using an exact solution based on the microscopic dynamics. For an arbitrary single-file system, we apply perturbation techniques and derive an explicit formula for the variance in terms of the transport coefficients. The same method also allows us to compute the fourth cumulant of the tagged particle displacement for the symmetric exclusion process.

**Keywords** Single-file diffusion, Macroscopic fluctuation theory, Anomalous diffusion, Large deviations

**PACS** 05.40.-a, 05.70.Ln, 05.70.Np, 05.10.Gg

## 1 Introduction

In non-equilibrium statistical mechanics, dynamical properties of interacting many-body systems play as important role as non-equilibrium steady states. Unlike the equilibrium properties, dynamical properties depend on the initial and boundary conditions and on details of the interactions.

The motion of an individual particle in a system of interacting particles is a fundamental *dynamical* problem in statistical mechanics even when the entire system is in equilibrium, or in a non-equilibrium steady state. A single Brownian particle exhibits normal diffusion—the mean-square displacement grows linearly with time. In a gas composed of impenetrable diffusing particles, the mobility of an individual particle is reduced due to the presence of other particles. In two and three dimensions, this effect can be significant in dense gases, but qualitatively the tagged particle still undergoes normal diffusion. In one dimension, however, the order of particles is preserved and transport can become anomalously slow. Transport in such systems is known as single-file diffusion.

Single-file diffusion is prevalent in numerous physical, chemical and biological processes: molecular motion inside porous medium like zeolite [1, 2], water transport inside a carbon nanotube [3], motion of tagged monomers in a polymer chain [4], sliding of proteins in a DNA sequence [5], ion channels through biological membranes [6], super-ionic conductors [7], etc.

The motion of an individual (tagged) particle strongly depends on whether the dynamics is biased or not. A diffusive behavior emerges when the dynamics is biased [8–14]. In the unbiased case, the

---

P L Krapivsky  
Physics Department, Boston University, Boston, Massachusetts 02215, USA  
Kirone Mallick and Tridib Sadhu  
Institut de Physique Théorique, CEA/Saclay, F-91191 Gif-sur-Yvette Cedex, France

variance of the displacement  $X_T$  exhibits a sub-diffusive  $\sqrt{T}$  growth. In this paper, we focus on the unbiased situation.<sup>1</sup>

The sub-diffusive  $\langle X_T^2 \rangle \sim \sqrt{T}$  behavior was first derived by Harris [17] for the system of Brownian particles interacting only through the hard-core repulsion; the same qualitative behavior was found to occur [18, 19] for the symmetric exclusion process (SEP). Since then, a large amount of work has been devoted to the study of more general single-file systems [20–30]. Most of these studies concentrate on the calculation of the variance, and even this requires a rather elaborate analysis when there are interactions in addition to the hard-core repulsion. The sub-diffusive scaling has also been verified in several experimental systems [3, 31–35].

Additional interest has been triggered by the connection to interface fluctuations. The problem of tagged particle displacement can be mapped to the height fluctuation of a one-dimensional interface [10]. The unbiased case is related to the Edwards-Wilkinson interface growth [36], whereas the biased case resembles the Kardar-Parisi-Zhang interface growth [37].

An analytical treatment of single-file diffusion is a challenging many-body problem due to the correlations caused by the non-crossing condition. The sub-diffusive scaling of the displacement of the tagged particle,  $\langle X_T^2 \rangle \sim \sqrt{T}$ , is easy to understand heuristically [7, 38–40]. An intriguing and a bit counter-intuitive property of single-file diffusion is that the variance depends on the initial state. More precisely, the scaling of the variance is robust,  $\langle X_T^2 \rangle \sim \sqrt{T}$ , but the pre-factor depends on whether we average only over the realizations of the stochastic dynamics or we additionally perform the averaging over initial positions [14, 41, 42].

Our goal is to study the higher cumulants, in fact the full statistics of the tagged particle displacement. The cumulants are encoded in the cumulant generating function

$$\mu_T(\lambda) = \log \langle e^{\lambda X_T} \rangle = \lambda \langle X_T \rangle_c + \frac{\lambda^2}{2!} \langle X_T^2 \rangle_c + \frac{\lambda^3}{3!} \langle X_T^3 \rangle_c + \dots, \quad (1)$$

where  $\langle X_T^n \rangle_c$  is the  $n^{\text{th}}$  cumulant and  $\lambda$  a fugacity parameter. For example, the first cumulant is the average and the second cumulant is the variance.

For impenetrable Brownian particles and for the SEP all the cumulants of the displacement scale as  $\sqrt{T}$ , see [19, 43]. The same is expected to hold for more general single-file systems, as we shall show explicitly using a hydrodynamic formulation. In long time limit, the probability of the rescaled position of the tagged particle has a large deviation form

$$P\left(\frac{X_T}{\sqrt{4T}} = x\right) \asymp e^{-\sqrt{4T} \phi(x)}. \quad (2)$$

Here  $\phi(x)$  is the large deviation function. The symbol  $\asymp$  implies that the logarithms exhibit the same asymptotic behavior:  $A \asymp B$  means that  $\lim_{T \rightarrow \infty} \frac{\log A}{\log B} = 1$ . The cumulant generating function  $\mu_T(\lambda)$  is related to  $\phi(x)$  via a Legendre transform [44, 45].

In the present work we apply the macroscopic fluctuation theory (MFT) to single-file diffusion. The MFT was developed by Bertini, De Sole, Gabrielli, Jona-Lasinio and Landim [46–50] for calculating large deviation functions in classical diffusive systems; similar results were obtained in the context of shot noise in conductors [51, 52]. The MFT provides a significant step towards constructing a general theoretical framework for non-equilibrium systems [44, 53]. Over the past decade the MFT has been successfully applied to numerous systems [46–48, 53–70]. A perfect agreement between the MFT and microscopic calculations has been observed whenever results from both approaches were available. The MFT is a powerful and versatile tool, although the analysis is involved and challenging in most cases.

The MFT allows one to probe large deviations of macroscopic quantities such as the total current. Intriguingly, large deviations of an individual microscopic particle in a single-file system can be captured by the MFT [29, 69]. The essential property of single-file systems—the fixed order of the particles—allows us to express the displacement of the tagged particle in a way amenable to the MFT treatment.

The MFT is an outgrowth of fluctuating hydrodynamic [71]. Using a path integral formulation one associates a classical action to a particular time evolution of the system. The optimal path has the least action and the analysis boils down to solving the Hamilton equations corresponding to the least-action paths. The advantage is that within this formulation all the microscopic details of the

<sup>1</sup> If the particle dynamics is ballistic, the tagged particle exhibits a normal diffusion behavior [15, 16].

system are embedded in terms of two transport coefficients: diffusivity  $D(\rho)$  and mobility  $\sigma(\rho)$ . These are bulk properties of the system near equilibrium, which for an arbitrary single-file system can either be measured in experiment or calculated from the microscopic dynamics.

We shall show how the cumulant generating function  $\mu_T(\lambda)$ , equivalently the large deviation function  $\phi(x)$ , can be formally expressed in terms of  $D(\rho)$  and  $\sigma(\rho)$ . The  $\sqrt{T}$  scaling of the cumulants comes out from this formal solution. The dependence of the cumulants on the initial state is also naturally incorporated within this formalism. Different initial states lead to different boundary conditions, while the governing Hamilton equations remain the same.

Mathematically, the Hamilton equations are a pair coupled non-linear partial differential equations for two scalar fields. In the general case when  $D(\rho)$  and  $\sigma(\rho)$  are arbitrary, the Hamilton equations are intractable. The only solvable case corresponds to Brownian particles where  $D(\rho) = 1$  and  $\sigma(\rho) = 2\rho$ . For this single-file system we deduce a closed form formula for the cumulant generating function and the associated large deviation function. We analyze annealed and quenched<sup>2</sup> initial states. Our results show that the single-file system remains forever sensitive to the initial state—the large-deviation function has a very different behavior in the two cases. We verify the MFT predictions in the particular case of Brownian particles by comparing with exact results which we derive using the microscopic dynamics: the large deviation functions coming from these independent methods match perfectly.

In the general case of arbitrary  $D(\rho)$  and  $\sigma(\rho)$  we use a series expansion method. In principle, the cumulants of all order can be evaluated iteratively, but in practice the calculations become very cumbersome as the order increases. Our analysis leads to an exact result for the variance of the tagged particle in terms of  $D(\rho)$  and  $\sigma(\rho)$ . The variance in the annealed and quenched cases differ by  $\sqrt{2}$ . This was observed for the symmetric random average process [72] and for impenetrable Brownian particles [42], and it remains generally valid for single-file diffusion. Our general formula applies to an arbitrary single-file system and it encompasses all the results derived for specific models [17, 18, 21, 42] as well as experimental results [33]. For the SEP, where the transport coefficients are  $D(\rho) = 1$  and  $\sigma(\rho) = 2\rho(1 - \rho)$ , we derive an explicit formula for the fourth cumulant. Our derivation is based on the macroscopic MFT framework, yet we also needed exact results for the integrated current which were derived in Ref. [73] on the basis of an exact *microscopic* analysis employing a Bethe ansatz.

The remainder of this paper is structured as follows. In Section 2 we start from a fluctuating hydrodynamics and show how the statistics of the tagged particle in single-file diffusion can be integrated into the MFT, leading to a variational problem. In Section 3, we analyze the single-file system of Brownian point particles. A perturbative treatment of the Hamilton equations with arbitrary  $D(\rho)$  and  $\sigma(\rho)$  is presented in Section 4. This allows us to derive a general formula for the variance of  $X_T$ . In Section 5, we calculate the fourth cumulant for the SEP. We present a microscopic analysis of the single-file system of Brownian particles in Section 6. Some intermediate technical steps are relegated to the Appendices.

## 2 A hydrodynamic formulation

The macroscopic fluctuation theory (MFT) is a deterministic re-formulation of fluctuating hydrodynamics. It has proven to be a very successful framework [50, 61] for probing large deviations in diffusive particle systems (such as lattice gases). The MFT generalizes the Freidlin–Wentzell theory [74] of finite-dimensional dynamical systems with random perturbations to a class of stochastic infinite-dimensional dynamical systems.

The simplicity of lattice gases, and generally diffusive particle systems, is that on a macroscopic level the full description is provided by one scalar field, the density, satisfying the diffusion equation. Taking into account the stochasticity of underlying microscopic dynamics leads to a more comprehensive description known as fluctuating hydrodynamics [71]. The basic assumption is that the behavior is still essentially hydrodynamic on a large (in comparison e.g. with the lattice spacing and the hopping time) length and time scales. In one dimension, the fluctuating hydrodynamics is based on the Langevin equation

$$\partial_t \rho = \partial_x \left[ D(\rho) \partial_x \rho + \sqrt{\sigma(\rho)} \, \eta \right], \quad (3)$$

<sup>2</sup> We shall often use this concise, but a bit imprecise, description. A quenched state is a fixed initial state; an annealed state is actually a collection of states, say states with macroscopically uniform density, and we perform averaging over all these states.

where  $\eta = \eta(x, t)$  is a Gaussian noise with covariance

$$\langle \eta(x, t) \eta(x', t') \rangle = \delta(x - x') \delta(t - t'). \quad (4)$$

The diffusion coefficient  $D(\rho)$  and the mobility  $\sigma(\rho)$  are related to the free energy density  $f(\rho)$  through the fluctuation-dissipation relation [44, 62, 71]

$$f''(\rho) = \frac{2D(\rho)}{\sigma(\rho)} \quad (5)$$

which is a consequence of the assumption of local equilibrium.

The key feature of the fluctuating hydrodynamics and the MFT is that all microscopic details (interactions, hopping rules, etc.) are embedded into the two transport coefficients  $D(\rho)$  and  $\sigma(\rho)$ . These transport coefficients generically depend on the density and can be difficult to determine analytically, but they can be measured numerically or experimentally.

The governing Langevin equation (3) describes long time properties of the system where the density varies smoothly over the coarse graining scale. The analysis based on this equation correctly captures the large scale statistics of the fluctuations. For example, in the case of single-file diffusion only the leading time-dependence of the cumulants is correctly captured by fluctuating hydrodynamics (equivalently, the MFT).

The remarkable property that allows one to apply the MFT to single-file diffusion is a simple relation between the position of the tagged particle and the density  $\rho(x, t)$  [29]. Let the tagged particle starts at the origin at time  $t = 0$  and moves within a time window  $[0, T]$ . Its position  $X_t$  at any time  $t$  is related to the density by the single-file constraint that particles do not cross each other. This gives

$$\int_{X_t[\rho]}^{\infty} dx \rho(x, t) = \int_0^{\infty} dx \rho(x, 0), \quad (6)$$

thereby defining the tagged particle position<sup>3</sup> as a functional  $X_t[\rho]$  of the density profile  $\rho(x, t)$ . Equivalently, we can write

$$\int_0^{X_t[\rho]} dx \rho(x, t) = \int_0^{\infty} dx (\rho(x, t) - \rho(x, 0)). \quad (7)$$

In contrast to (6), both integrals in (7) are convergent.

## 2.1 Variational formulation

The generating function of the cumulants of the tagged particle position  $X_T$  at time  $T$  can be written as a path integral of the density profile as

$$\langle e^{\lambda X_T} \rangle = \left\langle \int \mathcal{D}[\rho] e^{\lambda X_T[\rho]} \delta \left( \partial_t \rho - \partial_x [D(\rho) \partial_x \rho + \sqrt{\sigma(\rho)} \eta] \right) \right\rangle. \quad (8)$$

The Dirac delta function enforces the validity of the Langevin equation (3). The angular brackets denote the averaging that can consist of two parts: over the initial density field  $\rho(x, 0)$  and over realizations of the stochastic noise  $\eta(x, t)$  in the time window  $[0, T]$  (i.e., the averaging over the history of evolution of the density profile).

If we allow the initial state to fluctuate, we must include the probability  $\text{Prob}[\rho(x, 0)]$  of the initial density profile. It is useful to define the function

$$F[\rho(x, 0)] = -\log (\text{Prob}[\rho(x, 0)]). \quad (9)$$

For an initial state at equilibrium (annealed case),  $F[\rho]$  is related to the free energy. For a quenched initial state, there are no fluctuations of the initial profile at  $t = 0$  and we take the initial density profile to be the uniform profile with density  $\rho$ .

---

<sup>3</sup>  $X_t[\rho]$  is not uniquely fixed when there are regions with density equal to zero, but such configurations are highly improbable.

The following analysis is essentially the Martin-Siggia-Rose formalism [75, 76]. Incorporating the average over the initial state, the generating function becomes

$$\langle e^{\lambda X_T} \rangle = \int \mathcal{D}[\rho] e^{\lambda X_T[\rho] - F[\rho(x,0)]} \left\langle \delta \left( \partial_t \rho - \partial_x [D(\rho) \partial_x \rho + \sqrt{\sigma(\rho)} \eta] \right) \right\rangle_\eta.$$

The subscript  $\eta$  denotes the average over the history of noise within the window  $[0, T]$  and  $F$  appears only in the annealed case. Replacing the delta function by an integral over a field  $\hat{\rho}$  we get

$$\langle e^{\lambda X_T} \rangle = \int \mathcal{D}[\rho] e^{\lambda X_T[\rho] - F[\rho(x,0)]} \left\langle \int \mathcal{D}[\hat{\rho}] e^{-\int_0^T dt \int_{-\infty}^{\infty} dx \hat{\rho} (\partial_t \rho - \partial_x [D(\rho) \partial_x \rho + \sqrt{\sigma(\rho)} \eta])} \right\rangle_\eta.$$

The average with respect to the Gaussian variable  $\eta(x, t)$  is computed to yield

$$\langle e^{\lambda X_T} \rangle = \int \mathcal{D}[\rho, \hat{\rho}] e^{-S_T[\rho, \hat{\rho}]}$$

with action

$$S_T[\rho, \hat{\rho}] = -\lambda X_T[\rho] + F[\rho(x, 0)] + \int_0^T dt \int_{-\infty}^{\infty} dx \left( \hat{\rho} \partial_t \rho - H(\rho, \hat{\rho}) \right) \quad (10)$$

and Hamiltonian

$$H(\rho, \hat{\rho}) = \frac{\sigma(\rho)}{2} (\partial_x \hat{\rho})^2 - D(\rho) (\partial_x \rho) (\partial_x \hat{\rho}). \quad (11)$$

(In deriving these equations, we assumed that  $\hat{\rho}(x, t) \rightarrow 0$  when  $x \rightarrow \pm\infty$ .)

Let us rescale time by  $T$  and space by  $\sqrt{T}$ . The tagged particle position is also rescaled by  $\sqrt{T}$ , so the action  $S_T[\rho, \hat{\rho}]$  is proportional to  $\sqrt{T}$  and we can write  $S_T[\rho, \hat{\rho}] = \sqrt{T} S[p, q]$ . This simple observation leads to the anomalous scaling in single-file diffusion. Indeed, for large  $T$ , the path integral is dominated by the least action. The cumulant generating function is then given by

$$\mu_T(\lambda) = -\sqrt{T} S[p, q], \quad (12)$$

where we denote by  $(p, q) \equiv (\hat{\rho}, \rho)$  the optimal paths of the least action.<sup>4</sup> Equation (12) implies that at large time all the cumulants of the tagged particle position scale as  $\sqrt{T}$ . Note that the analysis captures only the leading  $T$  dependence of the cumulants, and the sub-leading terms of the cumulants come from the correction to the saddle point approximation.

#### *Paths of least action*

To determine the paths of least action we consider a small variation around  $(p, q)$  as  $\rho = q + \delta\rho$  and  $\hat{\rho} = p + \delta\hat{\rho}$ . The variation of the action is then

$$\begin{aligned} \delta S_T &= \int_0^T dt \int_{-\infty}^{\infty} dx \left( \partial_t q - \frac{\delta H}{\delta p} \right) \delta \hat{\rho}(x, t) - \int_0^T dt \int_{-\infty}^{\infty} dx \left( \partial_t p + \frac{\delta H}{\delta q} \right) \delta \rho(x, t) \\ &+ \int_{-\infty}^{\infty} dx \left( -\lambda \frac{\delta X_T}{\delta q(x, 0)} + \frac{\delta F}{\delta q(x, 0)} - p(x, 0) \right) \delta \rho(x, 0) \\ &+ \int_{-\infty}^{\infty} dx \left( -\lambda \frac{\delta X_T}{\delta q(x, T)} + p(x, T) \right) \delta \rho(x, T). \end{aligned} \quad (13)$$

For the action to be stationary,  $\delta S_T[p, q] = 0$ , integrals in (13) must vanish. Since  $\delta \hat{\rho}(x, t)$  and  $\delta \rho(x, t)$  are arbitrary, vanishing of the integrals in the first line in (13) lead to the Hamilton equations

$$\partial_t q = \frac{\delta H}{\delta p} \quad \text{and} \quad \partial_t p = -\frac{\delta H}{\delta q}. \quad (14)$$

---

<sup>4</sup> The choice of notation,  $p$  and  $q$ , hints at the Hamiltonian nature of the governing equations (14).

The analysis of the boundary terms gives the boundary conditions for the optimal fields  $(p, q)$ . The density  $\delta\rho(x, T)$  at final time  $T$  is unconstrained, so the vanishing of the integral in the third line in (13) leads to

$$p(x, T) = \lambda \frac{\delta X_T}{\delta q(x, T)}. \quad (15)$$

The vanishing of the integral in the second line in (13) depends on the way the system is prepared. In the quenched case,  $\delta\rho(x, 0) = 0$  by definition so the integral vanishes. In the annealed setting,  $\delta\rho(x, 0)$  is arbitrary, therefore its pre-factor must vanish. Thus

$$q(x, 0) = \rho \quad \text{quenched case,} \quad (16)$$

$$p(x, 0) = -\lambda \frac{\delta X_T}{\delta q(x, 0)} + \frac{\delta F}{\delta q(x, 0)} \quad \text{annealed case.} \quad (17)$$

The first equation states that the initial density in the quenched case is assumed to be uniform.

For the rest of the paper, we denote the position of the tagged particle corresponding to the least action path by

$$Y \equiv X_T[q].$$

Note that  $X_T$  is in general a random variable depending on the history of the tagged particle, but  $Y$  is a deterministic quantity.

Overall, the stochastic problem of characterizing statistics of the tagged particle position reduces to solving a variation problem. We now write the Hamilton equations and the corresponding boundary conditions by computing the functional derivatives. Different boundary conditions emerge for the quenched and annealed settings.

#### *Quenched case*

The Hamilton equations read

$$\partial_t p + D(q) \partial_{xx} p = -\frac{\sigma'(q)}{2} (\partial_x p)^2 \quad (18)$$

$$\partial_t q - \partial_x (D(q) \partial_x q) = -\partial_x (\sigma(q) \partial_x p) \quad (19)$$

where  $\sigma'(q) = d\sigma(q)/dq$ . Using (6, 7) we compute the functional derivatives of  $Y$  (see Appendix A) and obtain the boundary conditions

$$q(x, 0) = \rho \quad \text{and} \quad p(x, T) = \frac{\lambda}{q(Y, T)} \Theta(x - Y), \quad (20)$$

where  $\Theta(x)$  is the Heaviside step function.

The corresponding minimal action (10) yields the cumulant generating function

$$\mu_{\mathcal{Q}}(\lambda) = \lambda Y - \int_0^T dt \int_{-\infty}^{\infty} dx \left[ p \partial_t q - \frac{\sigma(q)}{2} (\partial_x p)^2 + D(q) (\partial_x q) (\partial_x p) \right].$$

Hereinafter we use the subscript  $\mathcal{Q}$  to denote the quenched case. (For the sake of clarity, the subscript  $T$  denoting the time variable is omitted from the following formulas.) The cumulant generating function can be further simplified by using Eqs. (18)–(19) and integrating by parts. One gets

$$\mu_{\mathcal{Q}}(\lambda) = \lambda Y - \int_0^T dt \int_{-\infty}^{\infty} dx \frac{\sigma(q)}{2} (\partial_x p)^2. \quad (21)$$

### Annealed case

The Hamilton equations are the same as in the quenched case, Eqs. (18)–(19). The boundary conditions are different. To derive them we use the well-known relation (see [44,49] and [62] showing the consistency with the MFT and the fluctuation dissipation relation)

$$F[\rho(x, 0)] = \int_{-\infty}^{\infty} dx \int_{\rho}^{\rho(x, 0)} dr \frac{2D(r)}{\sigma(r)} (\rho(x, 0) - r). \quad (22)$$

Computing the functional derivative of  $Y$  and  $F[q(x, 0)]$ , we recast (15) and (17) into

$$p(x, T) = \frac{\lambda}{q(Y, T)} \Theta(x - Y), \quad (23)$$

$$p(x, 0) = \frac{\lambda}{q(Y, T)} \Theta(x) + \int_{\rho}^{q(x, 0)} dr \frac{2D(r)}{\sigma(r)}. \quad (24)$$

The least action (10) can be simplified using (22) and the Hamilton equations. The cumulant generating function for the annealed initial state becomes

$$\mu_{\mathcal{A}}(\lambda) = \lambda Y - \int_{-\infty}^{\infty} dx \int_{\rho}^{q(x, 0)} dr \frac{2D(r)}{\sigma(r)} (q(x, 0) - r) - \int_0^T dt \int_{-\infty}^{\infty} dx \frac{\sigma(q)}{2} (\partial_x p)^2. \quad (25)$$

Hereinafter the subscript  $\mathcal{A}$  denotes the annealed case.

*A Remark on Symmetry.* It is instructive to look at the symmetry properties of the action and optimal fields. One can notice symmetry relations

$$q_{-\lambda}(-x, t) = q_{\lambda}(x, t) \quad (26)$$

$$p_{-\lambda}(-x, t) = p_{\lambda}(x, t) - \frac{\lambda}{q_{\lambda}(Y_{\lambda}, T)} \quad (27)$$

$$Y_{-\lambda} = -Y_{\lambda} \quad (28)$$

$$\mu_T(-\lambda) = \mu_T(\lambda) \quad (29)$$

underlying the variational problem in both quenched and annealed settings. Since  $\mu(\lambda)$  is an even function, all odd cumulants of the tagged particle position are zero. This is the consequence of the fact that the microscopic dynamics is unbiased.

### 3 Brownian particles with hard-core repulsion

It is hard to analyze the variational problem for a general single-file system characterized by density-dependent transport coefficients  $D(\rho)$  and  $\sigma(\rho)$ . The only system that is amenable to analyses, both macroscopic (based on the MFT) and microscopic, is the system of Brownian point particles with hard-core repulsion (Fig. 1). The system was first studied by Harris [17] who used a mapping to non-interacting particles to compute the variance of the tagged particle position. The problem was later studied by many other authors (see e.g. [20, 25, 42, 77]).

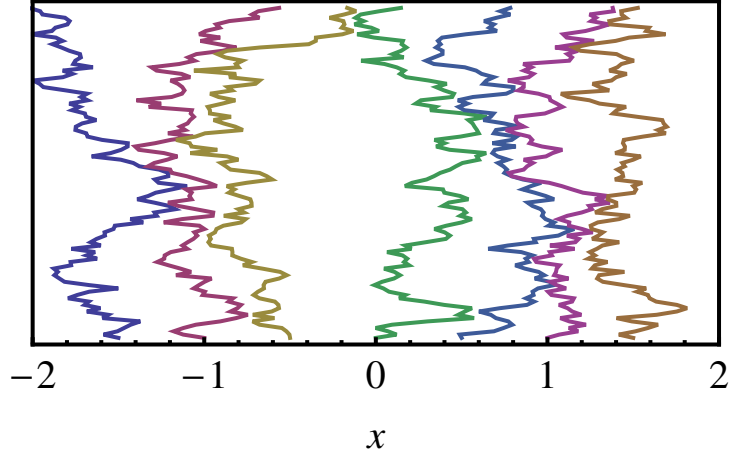
For Brownian particles the diffusion coefficient is constant,  $D(\rho) = 1$ , and the mobility is a linear function of density,  $\sigma(\rho) = 2\rho$ . The Hamilton equations (18)–(19) become

$$\partial_t p + \partial_{xx} p = -(\partial_x p)^2, \quad (30)$$

$$\partial_t q - \partial_{xx} q = -\partial_x (2q \partial_x p). \quad (31)$$

A canonical Hopf-Cole transformation,  $(p, q) \rightarrow (P, Q) = (e^p, qe^{-p})$ , is known [59, 65, 78] to simplify Eqs. (30)–(31). The new conjugate variables satisfy the Hamilton equations with  $H = -(\partial_x P)(\partial_x Q)$ . In the  $(P, Q)$  variables, the governing equations are (linear) anti-diffusion and diffusion equations:

$$\partial_t P + \partial_{xx} P = 0 \quad \text{and} \quad \partial_t Q - \partial_{xx} Q = 0. \quad (32)$$



**Fig. 1** A sample trajectory of Brownian point particles with hard-core repulsion. The trajectories may come infinitely close but never cross each other, keeping the order of particles unchanged.

Solving these equations and returning to the original variables  $(p, q)$  we arrive at a formal solution

$$p(x, t) = \log \left( \int_{-\infty}^{\infty} dz e^{p(z, T)} \frac{\exp \left( -\frac{(z-x)^2}{4(T-t)} \right)}{\sqrt{4\pi(T-t)}} \right), \quad (33)$$

$$q(x, t) = \int_{-\infty}^{\infty} dz q(z, 0) e^{p(x, t) - p(z, 0)} \frac{\exp \left( -\frac{(z-x)^2}{4t} \right)}{\sqrt{4\pi t}} \quad (34)$$

applicable to both quenched and annealed settings and to arbitrary  $p(x, T)$  and  $q(x, 0)$ .

### 3.1 Quenched case

Let us re-write (20) as

$$q(x, 0) = \rho \quad \text{and} \quad p(x, T) = B \Theta(x - Y). \quad (35)$$

Here we have used a shorthand notation

$$B = \frac{\lambda}{q(Y, T)}. \quad (36)$$

The quantity  $Y$  is related to  $q(x, t)$  via (7) which becomes

$$Y = \int_Y^{\infty} dx \left( \frac{q(x, T)}{\rho} - 1 \right). \quad (37)$$

in the quenched case.

The boundary conditions (35) depend on the solution itself. Let us proceed by treating  $B$  and  $Y$  as parameters to be determined later. The optimal fields in terms of  $B$  are

$$p(x, t) = \log \left[ 1 + (e^B - 1) \frac{1}{2} \operatorname{erfc} \left( \frac{Y - x}{\sqrt{4(T-t)}} \right) \right], \quad (38)$$

$$q(x, t) = \rho \int_{-\infty}^{\infty} dz \left( \frac{1 + (e^B - 1) \frac{1}{2} \operatorname{erfc} \left( \frac{Y - x}{\sqrt{4(T-t)}} \right)}{1 + (e^B - 1) \frac{1}{2} \operatorname{erfc} \left( \frac{Y - z}{\sqrt{4T}} \right)} \right) \frac{\exp \left( -\frac{(z-x)^2}{4t} \right)}{\sqrt{4\pi t}}, \quad (39)$$



where  $\text{erfc}(x)$  is the complementary error function. Substituting  $D(\rho) = 1$  and  $\sigma(\rho) = 2\rho$  into the formula for the cumulant generating function (21) leads to

$$\mu_{\mathcal{Q}}(\lambda) = \lambda Y - \int_0^T dt \int_{-\infty}^{\infty} dx q(\partial_x p)^2.$$

This expression for  $\mu_{\mathcal{Q}}(\lambda)$  can be simplified thanks to identity

$$q(\partial_x p)^2 = \partial_t(q p) - \partial_x(p \partial_x q - q \partial_x p - 2qp \partial_x p), \quad (40)$$

which results from Eqs. (30)–(31). Using (40) and  $p(x, t) \rightarrow 0$  as  $x \rightarrow \pm\infty$ , we obtain

$$\mu_{\mathcal{Q}}(\lambda) = \lambda Y - \int_{-\infty}^{\infty} dx q(x, T)p(x, T) + \int_{-\infty}^{\infty} dx q(x, 0)p(x, 0).$$

Plugging into this formula  $p(x, T)$  and  $q(x, 0)$  from (35) we get

$$\mu_{\mathcal{Q}}(\lambda) = \lambda Y - B \int_Y^{\infty} dx q(x, T) + \rho \int_{-\infty}^{\infty} dx p(x, 0).$$

It is clear from Eqs. (38)–(39) that the functions  $q(x, T)$  and  $p(x, 0)$  approach to non-zero constants as  $x \rightarrow \infty$ . This implies that the integrals in the above formula are not convergent although their linear combination is well defined. To write  $\mu_{\mathcal{Q}}(\lambda)$  in terms of convergent integrals we use the definition of  $Y$  in (37) and subsequently rearrange the integrals and obtain

$$\mu_{\mathcal{Q}}(\lambda) = (\lambda - \rho B) Y + \rho \int_Y^{\infty} dx (p(x, 0) - B) + \rho \int_{-Y}^{\infty} dx p(-x, 0).$$

The integrals are now written in terms of  $p(x, 0)$  and, using (38), the formula for the cumulant generating function becomes

$$\begin{aligned} \frac{\mu_{\mathcal{Q}}(\lambda)}{\sqrt{4T}} &= (\lambda - \rho B) y + \rho \int_0^{\infty} d\xi \log \left( 1 + \frac{e^{-B} - 1}{2} \text{erfc}(\xi) \right) \\ &\quad + \rho \int_0^{\infty} d\xi \log \left( 1 + \frac{e^B - 1}{2} \text{erfc}(\xi) \right). \end{aligned}$$

where  $y = Y/\sqrt{4T}$  and  $\xi = x/\sqrt{4T}$ . This is further simplified by using  $\text{erfc}(\xi) + \text{erfc}(-\xi) = 2$  to give

$$\frac{\mu_{\mathcal{Q}}(\lambda)}{\sqrt{4T}} = (\lambda - \rho B) y + \rho \int_0^{\infty} d\xi \log \left( 1 + \sinh^2 \left( \frac{B}{2} \right) \text{erfc}(\xi) \text{erfc}(-\xi) \right). \quad (41)$$

The reason for writing the formula in this form will become clear shortly.

So far,  $Y$  and  $B$  have been treated as parameters. One relation between these parameters is obtained by inserting (39) into (37):

$$y = \frac{e^B - 1}{2} \int_{-\infty}^{\infty} d\xi \frac{\text{erfc}(\xi - y) \text{erfc}(y - \xi)}{2 + (e^B - 1) \text{erfc}(y - \xi)}.$$

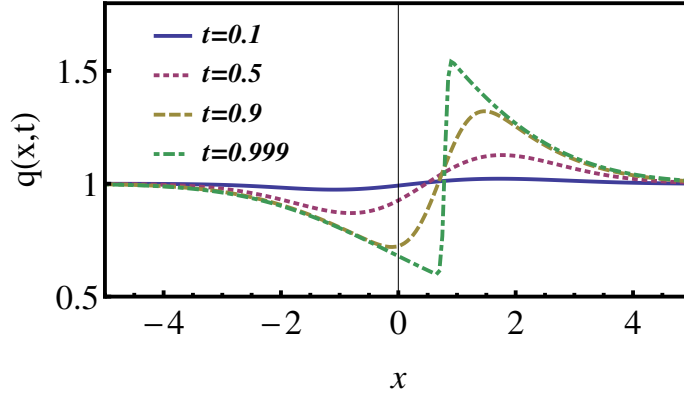
Changing the variable  $\xi \rightarrow \xi + y$  and using  $\text{erfc}(\xi) + \text{erfc}(-\xi) = 2$  we transform the above formula into

$$y = \frac{e^B - 1}{2} \int_0^{\infty} d\xi \frac{\text{erfc}(\xi) \text{erfc}(-\xi)}{2 + (e^B - 1) \text{erfc}(\xi)} - \frac{e^{-B} - 1}{2} \int_0^{\infty} d\xi \frac{\text{erfc}(\xi) \text{erfc}(-\xi)}{2 + (e^{-B} - 1) \text{erfc}(\xi)}$$

Massaging this formula one arrives at a more neat form

$$y = \frac{d}{dB} \int_0^{\infty} d\xi \log \left( 1 + \sinh^2 \left( \frac{B}{2} \right) \text{erfc}(\xi) \text{erfc}(-\xi) \right). \quad (42)$$

A similar self-consistent way of determining  $B$  using relation (36) does not lead to a unique value for  $B$ . This is because the solution  $q(x, t)$  in (39) is singular at  $(x, t) \equiv (Y, T)$ . A graphical representation



**Fig. 2** The optimum density field  $q(x, t)$ , Eq. (39), at different times. We set  $\rho = 1$ ,  $B = 2$  and  $T = 1$ , and determined  $Y$  from (42). The density at initial time  $t = 0$  starts from the quenched uniform profile  $q(x, 0) = 1$  and as time approaches  $t = T$ , the profile develops a sharp jump at the position of the tagged particle  $Y \simeq 1.54$ , indicating a discontinuity of the function  $q(x, T)$  at  $x = Y$ .

of this singularity is shown in Figure 2. The problem is analogous to the case of a diffusion equation with a step initial profile: The solution at any time  $t > 0$  is independent of the precise value of the initial profile at the position of the step.

For any value of  $B$ , Eqs. (38)–(39) give a solution of the Hamilton equations. Only one solution corresponds to the minimum action. This solution can be determined by optimizing the action with respect to  $B$ , i.e., by imposing

$$\frac{d\mu_{\mathcal{Q}}(\lambda)}{dB} = 0. \quad (43)$$

Combining this with (41) and (42) we find the optimal  $B$ :

$$B = \frac{\lambda}{\rho}. \quad (44)$$

Putting above in (41) we arrive at an explicit formula for the cumulant generating function:

$$\frac{\mu_{\mathcal{Q}}(\lambda)}{\sqrt{4T}} = \rho \int_0^\infty d\xi \log \left( 1 + \sinh^2 \left( \frac{\lambda}{2\rho} \right) \operatorname{erfc}(\xi) \operatorname{erfc}(-\xi) \right). \quad (45)$$

#### The large deviation function

The large deviation function is related to  $\mu_{\mathcal{Q}}(\lambda)$  via the Legendre transform:  $\phi_{\mathcal{Q}}(y) = \sup_{\lambda} \left( \lambda y - \frac{\mu_{\mathcal{Q}}(\lambda)}{\sqrt{4T}} \right)$ . Using (45) and (42) we can represent the large deviation function in the parametric form

$$\rho^{-1} \phi_{\mathcal{Q}}(y) = By - \int_0^\infty d\xi \log \left( 1 + \sinh^2 \left( \frac{B}{2} \right) \operatorname{erfc}(\xi) \operatorname{erfc}(-\xi) \right) \quad (46)$$

with  $B$  determined from the optimality requirement

$$\frac{d\phi_{\mathcal{Q}}}{dB} = 0. \quad (47)$$

An equivalent representation of  $\phi_{\mathcal{Q}}(y)$  is

$$\rho^{-1} \phi_{\mathcal{Q}}(y) = - \int_{-y}^\infty d\xi \log \left( 1 + \frac{e^{-B} - 1}{2} \operatorname{erfc}(\xi) \right) - \int_y^\infty d\xi \log \left( 1 + \frac{e^B - 1}{2} \operatorname{erfc}(\xi) \right). \quad (48)$$

All previous results have been derived using a macroscopic approach. In Section 6 we show that the same expression for  $\phi_{\mathcal{Q}}$  follows from an exact microscopic analysis. This is reassuring since the macroscopic approach is not fully rigorous, yet much more widely applicable than exact analyses which are limited to simplest systems.

### 3.2 Annealed case

The major difference with the quenched case comes from the boundary conditions (24)–(23). When  $D(\rho) = 1$  and  $\sigma(\rho) = 2\rho$ , the boundary conditions become

$$q(x, 0) = \rho \exp(p(x, 0) - B \Theta(x)), \quad (49)$$

$$p(x, T) = B \Theta(x - Y). \quad (50)$$

The parameter  $B$  is again defined in (36) and  $Y$  is determined from (7), equivalently

$$\int_0^Y dx q(x, T) = \int_0^\infty dx (q(x, T) - q(x, 0)). \quad (51)$$

The boundary condition depends on the solution itself and similar to the quenched case, a solution of the optimal fields is found by treating  $Y$  and  $B$  as parameters. They are later determined using the solution. Substituting the boundary condition in the general solution (33)–(34) leads to a formula of the optimal fields in terms of the  $Y$  and  $B$ .

$$p(x, t) = \log \left( 1 + (e^B - 1) \frac{1}{2} \operatorname{erfc} \left( \frac{Y - x}{\sqrt{4(T - t)}} \right) \right), \quad (52)$$

$$\frac{q(x, t)}{\rho} = \left( 1 + (e^{-B} - 1) \frac{1}{2} \operatorname{erfc} \left( \frac{x - Y}{\sqrt{4(T - t)}} \right) \right) \left( 1 + (e^B - 1) \frac{1}{2} \operatorname{erfc} \left( \frac{x}{\sqrt{4t}} \right) \right). \quad (53)$$

Equation (25) for the cumulant generating function becomes

$$\mu_{\mathcal{A}}(\lambda) = \lambda Y - \int_{-\infty}^{\infty} dx \int_{\rho}^{q(x, 0)} dr \left( \frac{q(x, 0)}{r} - 1 \right) - \int_0^T dt \int_{-\infty}^{\infty} dx q(\partial_x p)^2,$$

where we have used again  $D(\rho) = 1$  and  $\sigma(\rho) = 2\rho$ . This formula for  $\mu_{\mathcal{A}}(\lambda)$  can be rewritten, thanks to the identity (40), as

$$\begin{aligned} \mu_{\mathcal{A}}(\lambda) = & \lambda Y - \int_{-\infty}^{\infty} dx q(x, 0) \log \left( \frac{q(x, 0)}{\rho} \right) + \int_{-\infty}^{\infty} dx (q(x, 0) - \rho) \\ & - \int_{-\infty}^{\infty} dx q(x, T) p(x, T) + \int_{-\infty}^{\infty} dx q(x, 0) p(x, 0), \end{aligned}$$

(we have also taken into account that  $p(x, t) = 0$  at  $x \rightarrow \pm\infty$  at all time  $t$ ). Using (49)–(51) one can greatly simplify the above expression:

$$\mu_{\mathcal{A}}(\lambda) = \lambda Y + \int_{-\infty}^{\infty} dx (q(x, 0) - \rho).$$

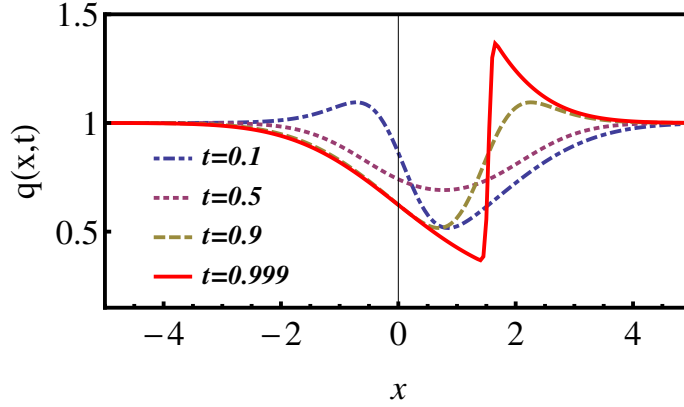
Combining this with (53) we find

$$\frac{\mu_{\mathcal{A}}(\lambda)}{\sqrt{4T}} = \lambda y + \rho \frac{e^B - 1}{2} \int_y^{\infty} d\xi \operatorname{erfc}(\xi) + \rho \frac{e^{-B} - 1}{2} \int_{-y}^{\infty} d\xi \operatorname{erfc}(\xi). \quad (54)$$

So far,  $Y$  and  $B$  were treated as parameters. Using (53) one can recast (51) into

$$2y = (e^B - 1) \int_y^{\infty} d\xi \operatorname{erfc}(\xi) - (e^{-B} - 1) \int_{-y}^{\infty} d\xi \operatorname{erfc}(\xi). \quad (55)$$

The second parameter  $B$  cannot be obtained by evaluating  $q(x, T)$  at  $x = Y$  because this function is singular. (This singularity is evident from Figure 3.) As in the quenched case, the parameter  $B$  has



**Fig. 3** The optimum density field  $q(x, t)$  in (53) for the annealed case at different times. We set  $Y = 1.5$  and  $T = 1$  and determined  $B$  from Eq. (55). As time approaches  $t = T = 1$ , the profile develops a sharp jump at the position  $Y$  of the tagged particle, indicating a discontinuity of the function  $q(x, T)$ . Note that the profile  $q(x, t)$  is symmetric under  $(x, t) \rightarrow (Y - x, T - t)$  as can be seen from Eq. (53).

to be determined from an optimization criterion  $\frac{d\mu_A}{dB} = 0$ . Using this together with (54) we derive a relation between  $B$ ,  $Y$  and the fugacity parameter  $\lambda$ :

$$\frac{2\lambda}{\rho} = (e^B - 1) \operatorname{erfc}(y) - (e^{-B} - 1) \operatorname{erfc}(-y). \quad (56)$$

To derive (56) we used

$$e^B \int_y^\infty d\xi \operatorname{erfc}(\xi) = e^{-B} \int_{-y}^\infty d\xi \operatorname{erfc}(\xi) \quad (57)$$

which follows from (55) in conjunction with the elementary identity

$$\int_{-y}^\infty d\xi \operatorname{erfc}(\xi) - \int_y^\infty d\xi \operatorname{erfc}(\xi) = 2y. \quad (58)$$

Equations (54)–(56) constitute a parametric solution for the cumulant generating function in the annealed setup. Unlike the quenched case, we don't have an explicit formula for  $\mu_A(\lambda)$ . Another parametric representation

$$\begin{aligned} \frac{\mu_A(\lambda)}{\rho\sqrt{4T}} &= \left( \frac{\lambda}{\rho} + \frac{1 - e^B}{1 + e^B} \right) y, \\ e^{2B} &= 1 + 2y \left[ \int_y^\infty du \operatorname{erfc}(u) \right]^{-1}, \\ \frac{\lambda}{\rho} &= (1 - e^{-B}) \left( 1 + \frac{e^B - 1}{2} \operatorname{erfc}(y) \right) \end{aligned}$$

has been reported in Refs. [29, 77]. It can be obtained from (54)–(56) using (57) and (58).

#### *The large deviation function*

The large deviation function is again the Legendre transform:  $\phi_A(y) = \sup_\lambda \left( \lambda y - \frac{\mu_A(\lambda)}{\sqrt{4T}} \right)$ . Combining it with (54) we obtain

$$\frac{2\phi_A(y)}{\rho} = -(e^B - 1) \int_y^\infty d\xi \operatorname{erfc}(\xi) - (e^{-B} - 1) \int_{-y}^\infty d\xi \operatorname{erfc}(\xi).$$

Using Eq. (57) we eliminate the dependence on  $B$  and arrive at the following explicit formula for the large deviation function

$$\frac{2\phi_{\mathcal{A}}(y)}{\rho} = \left\{ \sqrt{\int_y^\infty d\xi \operatorname{erfc}(\xi)} - \sqrt{\int_{-y}^\infty d\xi \operatorname{erfc}(\xi)} \right\}^2. \quad (59)$$

This result will be re-derived in Section 6 using an exact microscopic analysis.

### 3.3 Comparing annealed and quenched settings

#### *The cumulants*

By definition (1), the cumulants are obtained by expanding the cumulant generating function. For the quenched initial condition this expansion of  $\mu_{\mathcal{Q}}(\lambda)$  is simple to generate from the explicit formula (45). We write below the first few cumulants of the tagged particle position,

$$\langle X_T^2 \rangle_c = \frac{1}{\rho} I_1 \sqrt{T}, \quad (60)$$

$$\langle X_T^4 \rangle_c = \frac{1}{\rho^3} \left( I_1 - \frac{3}{2} I_2 \right) \sqrt{T}, \quad (61)$$

$$\langle X_T^6 \rangle_c = \frac{1}{\rho^5} \left[ I_1 - \frac{15}{2} (I_2 - I_3) \right] \sqrt{T}, \quad (62)$$

where  $I_n = \int_0^\infty dz [\operatorname{erfc}(z) \operatorname{erfc}(-z)]^n$ . The first two integrals  $I_1$  and  $I_2$  are known [79] leading to

$$\langle X_T^2 \rangle_c = \frac{\sqrt{2}}{\rho\sqrt{\pi}} \sqrt{T}, \quad (63)$$

$$\langle X_T^4 \rangle_c = \frac{1}{\rho^3} \left( \frac{9}{\pi} \arctan\left(\frac{1}{2\sqrt{2}}\right) - 1 \right) \frac{2\sqrt{2}}{\sqrt{\pi}} \sqrt{T}. \quad (64)$$

For the annealed case, the generating function has a parametric form. To make a series expansion in powers of  $\lambda$ , we first use Eq. (55) to eliminate  $B$  from (54) and subsequently make an expansion of the  $\mu_{\mathcal{A}}(\lambda)$  in powers of  $y$ . Then we use (55) and (56) to make an expansion of  $y$  in terms of  $\lambda$  leading to an expansion of  $\mu_{\mathcal{A}}(\lambda)$  in powers of  $\lambda$ . The first three non-trivial cumulants are:

$$\langle X_T^2 \rangle_c = \frac{2}{\rho\sqrt{\pi}} \sqrt{T}, \quad (65)$$

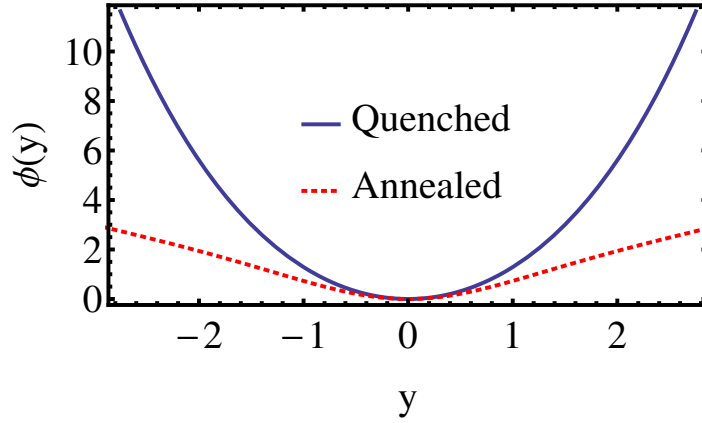
$$\langle X_T^4 \rangle_c = \frac{1}{\rho^3} \left( \frac{4}{\pi} - 1 \right) \frac{6}{\sqrt{\pi}} \sqrt{T}, \quad (66)$$

$$\langle X_T^6 \rangle_c = \frac{1}{\rho^5} \left( \frac{408}{\pi^2} - \frac{180}{\pi} + 18 \right) \frac{5}{\sqrt{\pi}} \sqrt{T}. \quad (67)$$

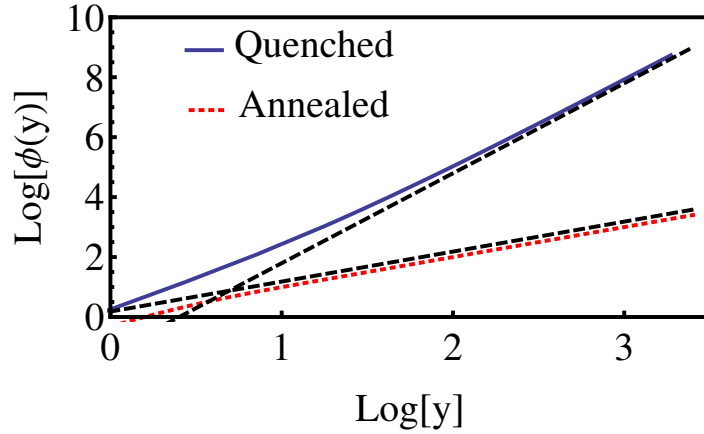
#### *The large deviation function*

The large deviation function depends on the setting. Strikingly different asymptotic behaviors emerge in the annealed and quenched settings (Figs. 4 and 5). In both cases, the large deviation function has a non-Gaussian tail. More precisely, the large deviation function  $\phi(y)$  grows linearly in the annealed case and cubically in the quenched case (see Fig. 5):

$$\phi_{\mathcal{A}}(y) \simeq \rho |y|, \quad \phi_{\mathcal{Q}}(y) \simeq \frac{\rho}{3} |y|^3. \quad (68)$$



**Fig. 4** The large deviation function  $\phi(y)$  of the tagged particle position in the annealed and quenched settings. In both cases, the initial average density is chosen to be uniform  $\rho = 1$ . Note that the quenched large deviation function is always larger, see Appendix B for a theoretical explanation.



**Fig. 5** Asymptotics of the large deviation functions for large values of  $y$ . The dashed straight lines correspond to the power laws  $y^3$  and  $y$ , for the quenched and the annealed case, respectively.

In the annealed case, the large  $y$  is derived from (59) and the asymptotic relations:

$$\begin{aligned} \int_y^\infty dz \frac{1}{2} \operatorname{erfc}(z) &= e^{-y^2} \left( \frac{1}{4\sqrt{\pi} y^2} + \mathcal{O}(y^{-4}) \right), \\ \int_{-y}^\infty dz \frac{1}{2} \operatorname{erfc}(z) &= y + e^{-y^2} \left( \frac{1}{4\sqrt{\pi} y^2} + \mathcal{O}(y^{-4}) \right). \end{aligned}$$

In the quenched case, the large deviation function has a parametric form, making the analysis a bit more involved. We use (46) and (47) and notice that large  $y$  corresponds to large  $B$ . The integral in (46) grows as  $\frac{2}{3}B^{3/2}$  for large  $B$ . Indeed, we divide the range of integration,  $\int_0^\infty dx = \int_0^{\sqrt{B}} dx + \int_{\sqrt{B}}^\infty dx$ , use the  $x \rightarrow \infty$  asymptotic

$$\operatorname{erfc}(x) \operatorname{erfc}(-x) = e^{-x^2} \left( \frac{2}{\sqrt{\pi} x} + \mathcal{O}\left(\frac{1}{x^3}\right) \right),$$

and write

$$\int_0^{\sqrt{B}} dx \log \left( 1 + \sinh^2 \left( \frac{B}{2} \right) \operatorname{erfc}(x) \operatorname{erfc}(-x) \right) = \int_0^{\sqrt{B}} dx (B - x^2) + \mathcal{O}(\sqrt{B}),$$

and find  $\frac{2}{3}B^{3/2}$ . This is the dominant contribution. The second integral contributes terms of order  $1/B$  only. Thus  $\rho^{-1}\phi_Q(y) \simeq By - \frac{2}{3}B^{3/2}$  which in conjunction with (47) leads to the asymptotic in (68).

A similar asymptotic dependence is also observed in the large deviation function for the current in the symmetric exclusion process with step initial condition [59, 65]. It would be interesting to study how universal are these power-law tails. An analogy between the displacement of the tagged particle and current suggests that the tails are non-universal, as for the extreme current fluctuations [63, 64].

#### 4 Calculation of the variance for general single-file systems

For a single-file system with arbitrary  $D(\rho)$  and  $\sigma(\rho)$  it is impossible to solve the Hamilton equations. One can try to seek an asymptotic solution [62] as a series expansion in powers of  $\lambda$ . This series solution is then used to expand the cumulant generating function.

The expansions for the optimal fields read

$$q(x, t) = \rho + \lambda q_1(x, t) + \lambda^2 q_2(x, t) + \dots, \quad (69)$$

$$p(x, t) = \lambda p_1(x, t) + \lambda^2 p_2(x, t) + \dots, \quad (70)$$

In the zeroth order, the solution is deterministic:  $q(x, t) = \rho$  and  $p(x, t) = 0$ . This follows from the governing Eqs. (18)–(19) both in the quenched and annealed cases. In the first order

$$\partial_t p_1 + D(\rho) \partial_{xx} p_1 = 0, \quad (71)$$

$$\partial_t q_1 - D(\rho) \partial_{xx} q_1 = -\sigma(\rho) \partial_{xx} p_1. \quad (72)$$

Using (21) and (25) we see that the variance

$$\frac{1}{2!} \langle X_T^2 \rangle = Y_1 - F_2 - \frac{\sigma(\rho)}{2} \int_0^T dt \int_{-\infty}^{\infty} dx (\partial_x p_1)^2 \quad (73)$$

explicitly depends on  $p_1, Y_1, F_2$ ; it also on  $q_1$  (through  $Y_1, F_2$ ). Since  $Y$  is an odd function of  $\lambda$ , see (28), the zeroth order term vanishes while the first order term is given by

$$Y_1 = \frac{1}{\rho} \int_0^{\infty} dx (q_1(x, T) - q_1(x, 0)) \quad (74)$$

as it follows from (6,7). It remains to determine  $F_2$ . In the quenched case,  $F$  does not appear (one can formally set  $F \equiv 0$  in the quenched setting). In the annealed case we use (22) and find

$$F_2 = \frac{D(\rho)}{\sigma(\rho)} \int_{-\infty}^{\infty} dx (q_1(x, 0))^2. \quad (75)$$

We now determine the variance  $\langle X_T^2 \rangle$  by solving (71)–(72) with the specific boundary conditions for the quenched and the annealed initial states.

##### 4.1 The quenched initial state

Plugging the series expansion into the boundary condition (20) yields

$$p_1(x, T) = \rho^{-1} \Theta(x) \quad \text{and} \quad q_1(x, 0) = 0 \quad (76)$$

in the linear order. With the above boundary condition the solution of (71) can be expressed as

$$\partial_x p_1(x, t) = \rho^{-1} g(0, T|x, t), \quad (77)$$

where  $g$  is the diffusion propagator

$$g(z, T|x, t) = \frac{1}{\sqrt{4\pi D(\rho)(T-t)}} \exp \left[ -\frac{(z-x)^2}{4D(\rho)(T-t)} \right] \quad (78)$$

for  $0 \leq t \leq T$ . Since  $p_1(x, t) = 0$  at  $x \rightarrow -\infty$ , we obtain

$$p_1(x, t) = \frac{1}{\rho} \int_{-\infty}^x dz g(0, T|z, t) = \frac{1}{2\rho} \operatorname{erfc} \left( \frac{-x}{2\sqrt{D(\rho)(T-t)}} \right). \quad (79)$$

Taking into account (77), the solution for  $q_1(x, t)$  in (72) can be written as  $q_1(x, t) = -\partial_x \psi(x, t)$  with

$$\psi(x, t) = \frac{\sigma(\rho)}{\rho} \int_0^t d\tau \int_{-\infty}^{\infty} dz g(x, t|z, \tau) g(0, T|z, \tau). \quad (80)$$

These are the only two quantities required for simplifying  $\langle X_T^2 \rangle$  given by Eq. (73). Recalling that  $q_1(x, 0) = 0$  (and that formally  $F_2 \equiv 0$  in the quenched initial state) we obtain

$$\frac{1}{2!} \langle X_T^2 \rangle_{\mathcal{Q}} = \frac{1}{\rho} \int_0^{\infty} dx q_1(x, T) - \frac{\sigma(\rho)}{2} \int_0^T dt \int_{-\infty}^{\infty} dx (\partial_x p_1)^2. \quad (81)$$

Using the solution for  $q_1(x, t)$ , the first integral yields

$$\frac{1}{\rho} \int_0^{\infty} dx q_1(x, T) = \frac{\psi(0, T)}{\rho}, \quad (82)$$

where we have taken into account that  $\psi(x, T)$  vanishes at  $x \rightarrow \infty$ . The second integral in (81) becomes

$$\frac{\sigma(\rho)}{2} \int_0^T dt \int_{-\infty}^{\infty} dx (\partial_x p_1)^2 = \frac{\psi(0, T)}{2\rho}. \quad (83)$$

Combining these two results we reduce the expression for the variance to

$$\frac{1}{2!} \langle X_T^2 \rangle_{\mathcal{Q}} = \frac{\psi(0, T)}{2\rho} = \frac{\sigma(\rho)}{2\rho^2} \int_0^T dt \int_{-\infty}^{\infty} dx [(g(0, T|x, t))]^2.$$

We use (78) and compute the integral. The final expression for the variance,

$$\langle X_T^2 \rangle_{\mathcal{Q}} = \frac{\sigma(\rho)}{\rho^2} \frac{\sqrt{T}}{\sqrt{2\pi D(\rho)}}, \quad (84)$$

is valid in the quenched setting in the general case of arbitrary  $D(\rho)$  and  $\sigma(\rho)$ .

#### 4.2 The annealed initial state

Inserting the series expansion (69)–(70) into the boundary condition (24)–(23) we obtain the boundary conditions in the first order

$$p_1(x, T) = \frac{\Theta(x)}{\rho} \quad \text{and} \quad q_1(x, 0) = \frac{\sigma(\rho)}{2D(\rho)} (p_1(x, 0) - p_1(x, T)). \quad (85)$$

The variance (73) becomes

$$\begin{aligned} \langle X_T^2 \rangle_{\mathcal{A}} = & \frac{1}{\rho} \int_0^{\infty} (q_1(x, T) - q_1(x, 0)) dx - \frac{D(\rho)}{\sigma(\rho)} \int_{-\infty}^{\infty} dx (q_1(x, 0))^2 \\ & - \frac{\sigma(\rho)}{2} \int_0^T dt \int_{-\infty}^{\infty} dx (\partial_x p_1(x, t))^2, \end{aligned} \quad (86)$$

where we have used (74) and (75).

The following the analysis is interwoven with the quenched case in a way that allows us to avoid the explicit computation of most integrals. First we note that in both cases the equation for  $p_1(x, t)$  and the boundary condition on  $p_1(x, T)$  are identical. Therefore (79) remains valid. The equation (72)



for  $q_1(x, t)$  is also the same in both cases, but the boundary condition (85) is different. The governing equation (72) is linear in  $q_1(x, t)$ , so we can write the solution as a sum

$$q_1(x, t) = q_i(x, t) + q_h(x, t), \quad (87)$$

where  $q_i(x, t)$  is the solution of the inhomogeneous equation

$$\partial_t q_i - D(\rho) \partial_{xx} q_i = -\sigma(\rho) \partial_{xx} p_1 \quad \text{with} \quad q_i(x, 0) = 0, \quad (88)$$

and  $q_h(x, t)$  is the solution of the homogeneous equation

$$\partial_t q_h - D(\rho) \partial_{xx} q_h = 0 \quad \text{with} \quad q_h(x, 0) = \frac{\sigma(\rho)}{2D(\rho)} (p_1(x, 0) - p_1(x, T)). \quad (89)$$

Comparing with the quenched case we notice that  $q_i(x, t)$  is same as  $q_1(x, t)$  in the quenched case:  $q_i(x, t) = -\partial_x \psi(x, t)$  with  $\psi(x, t)$  given by (80). Further, using (81) and (86) we find that the variance for the annealed case is related to the variance in the quenched case via

$$\frac{1}{2!} \langle X_T^2 \rangle_{\mathcal{A}} - \frac{1}{2!} \langle X_T^2 \rangle_{\mathcal{Q}} = -\frac{D(\rho)}{\sigma(\rho)} \int_{-\infty}^{\infty} dx (q_h(x, 0))^2 + \frac{1}{\rho} \int_0^{\infty} dx (q_h(x, T) - q_h(x, 0)). \quad (90)$$

The difference depends only on the homogeneous solution  $q_h(x, t)$ .

The second term on the right-hand side of Eq. (90) can be simplified using an identity

$$\frac{1}{\rho} \int_0^{\infty} dx (q_h(x, T) - q_h(x, 0)) = \frac{2D(\rho)}{\sigma(\rho)} \int_{-\infty}^{\infty} dx (q_h(x, 0))^2. \quad (91)$$

This identity is proved by using (89), then noting that  $\int_{-\infty}^{\infty} dx q_h(x, t) p_1(x, t)$  does not depend on time because  $p_1$  and  $q_h$  satisfy adjoint equations, and finally that recalling  $p_1(x, T) = \Theta(x)/\rho$ :

$$\begin{aligned} \frac{2D(\rho)}{\sigma(\rho)} \int_{-\infty}^{\infty} dx (q_h(x, 0))^2 &= \int_{-\infty}^{\infty} dx q_h(x, 0) (p_1(x, 0) - p_1(x, T)) \\ &= \int_{-\infty}^{\infty} dx p_1(x, T) (q_h(x, T) - q_h(x, 0)). \end{aligned}$$

Thus

$$\begin{aligned} \frac{1}{2!} \langle X_T^2 \rangle_{\mathcal{A}} - \frac{1}{2!} \langle X_T^2 \rangle_{\mathcal{Q}} &= \frac{D(\rho)}{\sigma(\rho)} \int_{-\infty}^{\infty} dx (q_h(x, 0))^2 \\ &= \frac{\sigma(\rho)}{4D(\rho)} \int_{-\infty}^{\infty} dx (p_1(x, T) - p_1(x, 0))^2, \end{aligned} \quad (92)$$

where in the last step we have used the boundary condition from Eq. (89). The last integral is computed using (79) to yield<sup>5</sup>

$$\frac{1}{2!} \langle X_T^2 \rangle_{\mathcal{A}} - \frac{1}{2!} \langle X_T^2 \rangle_{\mathcal{Q}} = \frac{\sigma(\rho)}{2\rho^2} \left( \frac{\sqrt{2}-1}{\sqrt{2\pi}} \right) \sqrt{\frac{T}{D(\rho)}}. \quad (93)$$

Combining this with (84) we establish a very simple general relation

$$\langle X_T^2 \rangle_{\mathcal{A}} = \sqrt{2} \langle X_T^2 \rangle_{\mathcal{Q}} \quad (94)$$

between the two initial states. This relation and Eq. (84) are applicable for any single-file diffusion, i.e., for arbitrary transport coefficients.<sup>6</sup> To verify the range of applicability of Eq. (84) let us compare with earlier work. In the simplest case of Brownian particles we recover the well-known expression for the variance (for more about Brownian particles see Refs. [17, 20, 25, 42]). For the SEP we also reproduce

<sup>5</sup> The difference  $\langle X_T^2 \rangle_{\mathcal{A}} - \langle X_T^2 \rangle_{\mathcal{Q}}$  is positive, as shown in Appendix B.

<sup>6</sup> Our derivations are based on the MFT which assumes local equilibrium. This is expected to be correct for systems with short-range inter-particle interactions.

the well-known result [18]. Let us also compare with the variance in single-file colloidal systems studied in Ref. [21]. For comparison, we need the relation

$$\sigma(\rho) = 2S(0, 0)D(\rho), \quad (95)$$

where  $S(0, 0)$  is the structure factor [71]. With this the variance for the annealed case becomes

$$\langle X_T^2 \rangle_{\mathcal{A}} = \frac{2S(0, 0)}{\rho^2} \sqrt{\frac{D(\rho)}{\pi}} \sqrt{T}, \quad (96)$$

which is identical to the one presented in [21]. For experimental measurements, it is convenient to express the variance in terms of the isothermal compressibility  $\kappa$ . Since  $S(0, 0) = \beta^{-1} \kappa \rho^2$ ,

$$\langle X_T^2 \rangle_{\mathcal{A}} = 2 \beta^{-1} \kappa \sqrt{\frac{D(\rho)}{\pi}} \sqrt{T}, \quad (97)$$

where  $\beta$  is the inverse temperature.

## 5 Fourth cumulant of a tagged particle in the SEP

The symmetric exclusion process (SEP) is a diffusive lattice gas. Each site is occupied by at most one particle and each particle attempts to hop to neighboring empty sites with unit rate. For the SEP, the diffusion coefficient is constant,  $D(\rho) = 1$ , while the mobility has the symmetric form  $\sigma(\rho) = 2\rho(1 - \rho)$  reflecting the  $\rho \leftrightarrow 1 - \rho$  mirror symmetry of the SEP. The formula for  $\sigma(\rho)$  can be derived using the fluctuation-dissipation relation (see Appendix C).

We employ a perturbative approach and compute the fourth cumulant of a tagged particle position. It proves useful to consider a one-parameter class of models with  $D(\rho) = 1$  and  $\sigma(\rho) = 2\rho(1 - \alpha\rho)$ . The parameter  $\alpha$  varies in the range  $0 \leq \alpha \leq 1$ , so the mobility remains positive. The case of  $\alpha = 0$  describes the single-file system of Brownian particles, whereas  $\alpha = 1$  corresponds to the SEP.

Mathematically, we want to solve equations (18)–(19). The corresponding boundary conditions depend on the solution itself. As discussed in Section 3, a way to solve the equations is by treating  $B = \lambda/q(Y, T)$  as a parameter. This determines the optimal fields in terms of  $B$  which we denote by  $q_B(x, t)$  and  $p_B(x, t)$ . The solution does not explicitly involve the fugacity parameter  $\lambda$  which only appears through the dependence of  $B$  on  $\lambda$ . A straightforward implementation of the definition of  $B$  does not lead to a unique value, as the function  $q(x, T)$  is singular at  $(x, t) \equiv (Y, T)$ . The value corresponding to the least action path is obtained by an optimization condition  $\frac{d\mu}{dB} = 0$ . This expresses  $B$  in terms of  $\lambda$  and leads to a parametric formula of the cumulant generating function

$$\mu = \lambda Y(B) - R(B) \quad \text{with} \quad \lambda \frac{dY(B)}{dB} = \frac{dR(B)}{dB}. \quad (98)$$

The function  $R(B)$  is defined in terms of the optimal fields  $q_B(x, t)$  and  $p_B(x, t)$ , Eq. (21) in the quenched case and Eq. (25) in the annealed case. A detailed implementation of this procedure has been presented in Section 3 for the system of Brownian particles where an exact solution was possible as the corresponding  $\sigma(\rho)$  is linear.

To proceed with the analysis for the quadratic  $\sigma(\rho)$ , it is instructive to recall the symmetry properties of the optimal fields (26)–(29):

$$q_{-B}(-x, t) = q_B(x, t) \quad \text{and} \quad p_{-B}(-x, t) = p_B(x, t) - B,$$

whereas  $Y(B)$  and  $\lambda(B)$  are odd functions of  $B$ . Combining all together shows that  $R(B)$  is an even function of  $B$ . This is consistent with the fact that the dynamics is unbiased.

It proves convenient to use  $B$  as a primary expansion parameter; at the end one can re-expand the results in terms of  $\lambda$  and compute the cumulants. The aforementioned symmetry properties allow us to seek  $Y(B)$ ,  $R(B)$  and  $\lambda(B)$  as the following expansions

$$\begin{aligned} Y(B) &= Y_1 B + Y_3 B^3 + \mathcal{O}(B^5) \\ R(B) &= R_2 B^2 + R_4 B^4 + \mathcal{O}(B^6) \\ \lambda(B) &= \lambda_1 B + \lambda_3 B^3 + \mathcal{O}(B^5). \end{aligned} \quad (99)$$

The last formula is also equivalent to

$$B = \left( \frac{1}{\lambda_1} \right) \lambda - \left( \frac{\lambda_3}{\lambda_1^4} \right) \lambda^3 + \mathcal{O}(\lambda^5). \quad (100)$$

Substituting the series expansions in the first equation of (98) leads to a formula for the cumulant generating function in powers of  $B$ ,

$$\mu = (\lambda_1 Y_1 - R_2) B^2 + (\lambda_1 Y_3 + \lambda_3 Y_1 - R_4) B^4 + \dots$$

The second equation of (98), rewritten as  $\lambda = \frac{dR}{dB} \left( \frac{dY}{dB} \right)^{-1}$  leads to

$$\lambda_1 = \frac{2R_2}{Y_1} \quad \text{and} \quad \lambda_3 = \frac{4R_4 - 3\lambda_1 Y_3}{Y_1}. \quad (101)$$

To determine the cumulants we use the above in the series expansion (1) and obtain

$$\langle X_T^2 \rangle_c = \frac{Y_1}{\lambda_1}, \quad (102)$$

$$\frac{1}{4!} \langle X_T^4 \rangle_c = \frac{1}{\lambda_1^4} (-R_4 + \lambda_1 Y_3). \quad (103)$$

These formal expressions hold in both annealed and quenched cases.

The boundary condition for the optimal fields derived in (24) and (23) can be rewritten as

$$p_B(x, 0) = B \Theta(x) + \int_{\rho}^{q_B(x, 0)} dr \frac{2D(r)}{\sigma(r)} = B \Theta(x) + \ln \frac{q_B(x, 0)(1 - \alpha\rho)}{\rho(1 - \alpha q_B(x, 0))} \quad (104)$$

$$p_B(x, T) = B \Theta(x - Y), \quad (105)$$

where we used the definition of  $B$  in (36). To proceed, we write an expansion of the optimal fields in powers of  $B$  as

$$q_B = \rho + q_1 B + q_2 B^2 + q_3 B^3 + \dots, \quad (106)$$

$$p_B = p_1 B + p_2 B^2 + p_3 B^3 + \dots. \quad (107)$$

Note that the  $q_k(x, t)$  and  $p_k(x, t)$  in the above formulas are different from those in (69)–(70) which were obtained using expansions in powers of  $\lambda$ , and not  $B$ .

A straightforward computation of the solution to different orders is tedious, it involves difficult integrals. We circumvent this by drawing comparison with the  $\alpha = 0$  case where an exact solution is available. We first illustrate this trick by computing the solution in the linear order in  $B$ . This will give us the second cumulant which we can compare with already known results (which were derived in the previous section in the general setup). Then we shall compute the forth cumulant.

### *The second cumulant*

The governing equations in the first order in  $B$  are

$$\begin{aligned} (\partial_t + \partial_{xx}) p_1 &= 0, \\ (\partial_t - \partial_{xx}) q_1 &= -2\rho(1 - \alpha\rho) \partial_{xx} p_1. \end{aligned}$$

To determine the corresponding boundary conditions we use a formal expansion of the step function,

$$\Theta(x - Y) = \Theta(x) - \delta(x)Y + \frac{1}{2} \delta'(x)Y^2 + \dots. \quad (108)$$

The boundary conditions are

$$p_1(x, T) = \Theta(x) \quad \text{and} \quad q_1(x, 0) = \rho(1 - \alpha\rho) (p_1(x, 0) - p_1(x, T)). \quad (109)$$

The solutions to the governing equations are almost the same as in the Brownian case:

$$p_1(x, t) = \hat{p}_1(x, t) \quad \text{and} \quad q_1(x, t) = (1 - \alpha\rho) \hat{q}_1(x, t), \quad (110)$$

where the hat denotes the corresponding solutions for  $\alpha = 0$  case, *i.e.*, for Brownian particles. In the rest of this paper we shall follow the same notation.

To derive the second cumulant we need to determine  $Y_1$ ,  $\lambda_1$  and  $R_2$ . Combining the series expansion for  $Y$  with the definition (51) results in

$$Y_1 = \frac{1}{\rho} \int_0^\infty dx (q_1(x, T) - q_1(x, 0)). \quad (111)$$

Thanks to (110) we have

$$Y_1 = (1 - \alpha\rho) \hat{Y}_1. \quad (112)$$

To compute  $R_2$ , we start with the general formula (25) which in the present case becomes

$$R = \int_{-\infty}^\infty dx \int_{\rho}^{q(x,0)} dr \frac{1}{r(1 - \alpha r)} (q(x, 0) - r) + \int_0^T dt \int_{-\infty}^\infty dx q(1 - \alpha q) (\partial_x p)^2. \quad (113)$$

At second order we get

$$R_2 = \frac{1}{2\rho(1 - \alpha\rho)} \int_{-\infty}^\infty dx (q_1(x, 0))^2 + \rho(1 - \alpha\rho) \int_0^T dt \int_{-\infty}^\infty dx (\partial_x p_1)^2 \quad (114)$$

which in conjunction with (110) yields

$$R_2 = (1 - \alpha\rho) \hat{R}_2. \quad (115)$$

The above relations (112) and (115) also establish

$$\lambda_1 = \hat{\lambda}_1. \quad (116)$$

Using (112) and (116) we transform (102) into

$$\langle X_T^2 \rangle_{\mathcal{A}} = (1 - \alpha\rho) \langle \hat{X}_T^2 \rangle_{\mathcal{A}}. \quad (117)$$

The variance  $\langle \hat{X}_T^2 \rangle_{\mathcal{A}}$  for the Brownian case is derived in (65) which then leads to

$$\langle X_T^2 \rangle_{\mathcal{A}} = 2 \frac{1 - \alpha\rho}{\rho} \frac{\sqrt{T}}{\sqrt{\pi}}. \quad (118)$$

This expression agrees with the general formula for the variance derived earlier in Section 4.2, validating the approach we used here.

The variance can also be determined by directly solving for  $q_1$  and  $p_1$ . An extension of this approach to the fourth cumulant requires explicit solution of the optimal fields up to the third order which is a very tedious task. The alternative approach based on the mapping to  $\alpha = 0$  case, as demonstrated above for the variance, considerably simplifies computations, so we adopt this approach in the following derivation of the fourth cumulant. Before proceeding, we write explicit formulas  $\lambda_1$  and  $Y_1$ . Using (56) which describes the case of  $\alpha = 0$ , we compute  $\hat{\lambda}_1 = \rho$  leading [due to (116)] to

$$\lambda_1 = \rho. \quad (119)$$

Similarly we derive

$$Y_1 = (1 - \alpha\rho) \frac{2}{\sqrt{\pi}} \sqrt{T}. \quad (120)$$

---

*The fourth cumulant*

In the second order in  $B$  we have

$$\begin{aligned}(\partial_t + \partial_{xx})p_2 &= -(1 - 2\alpha\rho)(\partial_x p_1)^2, \\ (\partial_t - \partial_{xx})q_2 &= -2\rho(1 - \alpha\rho)\partial_{xx}p_2 - 2(1 - 2\alpha\rho)\partial_x(q_1\partial_x p_1),\end{aligned}$$

whereas the corresponding equations for the third order are

$$\begin{aligned}(\partial_t + \partial_{xx})p_3 &= -2(1 - 2\alpha\rho)(\partial_x p_1)(\partial_x p_2) + 2\alpha q_1(\partial_x p_1)^2, \\ (\partial_t - \partial_{xx})q_3 &= -\partial_x \left[ 2\rho(1 - \alpha\rho)\partial_{xx}p_3 + 2(1 - 2\alpha\rho)q_1\partial_x p_2 + 2((1 - 2\alpha\rho)q_2 - \alpha(q_1)^2)\partial_x p_1 \right].\end{aligned}$$

These equation are derived from the Hamilton equations (18)–(19). The corresponding boundary conditions follow from (104)–(105):

$$\begin{aligned}p_2(x, T) &= -Y_1 \delta(x), \\ q_2(x, 0) &= \rho(1 - \alpha\rho)p_2(x, 0) + \frac{\rho(1 - \alpha\rho)(1 - 2\alpha\rho)}{2} [p_1(x, T) - p_1(x, 0)]^2, \\ p_3(x, T) &= \frac{1}{2}Y_1^2 \delta'(x), \\ q_3(x, 0) &= \rho(1 - \alpha\rho)p_3(x, 0) + \frac{1 - 2\alpha\rho}{\rho(1 - \alpha\rho)} q_1(x, 0)q_2(x, 0) - \frac{1 - 3\alpha\rho(1 - \alpha\rho)}{3[\rho(1 - \alpha\rho)]^2} [q_1(x, 0)]^3.\end{aligned}$$

From the above formulas, one can verify that the solutions at the second order are related to the corresponding solutions for  $\alpha = 0$  case by a simple transformation:

$$p_2(x, t) = (1 - 2\alpha\rho) \hat{p}_2(x, t) \Big|_{\hat{Y}_1 \rightarrow \frac{Y_1}{(1-2\alpha\rho)}}, \quad (121)$$

$$q_2(x, t) = (1 - \alpha\rho)(1 - 2\alpha\rho) \hat{q}_2(x, t) \Big|_{\hat{Y}_1 \rightarrow \frac{Y_1}{(1-2\alpha\rho)}}. \quad (122)$$

On the right-hand side we write  $\hat{Y}_1 \rightarrow \frac{Y_1}{(1-2\alpha\rho)}$  implying that  $\hat{Y}_1$  should be replaced by  $\frac{Y_1}{(1-2\alpha\rho)}$  in the corresponding solution for  $\alpha = 0$  case obtained by treating  $\hat{Y}_1$  as parameter.

A similar relation can be derived for the third order terms:

$$p_3(x, t) = (1 - 2\alpha\rho)^2 \hat{p}_3(x, t) \Big|_{\hat{Y}_1 \rightarrow \frac{Y_1}{(1-2\alpha\rho)}} + \alpha(1 - \alpha\rho)u(x, t), \quad (123)$$

$$q_3(x, t) = (1 - \alpha\rho)(1 - 2\alpha\rho)^2 \hat{q}_3(x, t) \Big|_{\hat{Y}_1 \rightarrow \frac{Y_1}{(1-2\alpha\rho)}} + \alpha(1 - \alpha\rho)^2 h(x, t). \quad (124)$$

Here  $u(x, t)$  is the solution of

$$(\partial_t + \partial_{xx})u = 2 \hat{q}_1 (\partial_x \hat{p}_1) \quad (125)$$

subject to  $u(x, T) = 0$ , while  $h(x, t)$  is the solution of

$$(\partial_t - \partial_{xx})h = -2 \rho \partial_{xx}u + 2 \partial_x [\hat{q}_1^2 \partial_x \hat{p}_1] \quad (126)$$

subject to  $h(x, 0) = \rho u(x, 0) - \frac{1}{3\rho} [\hat{q}_1(x, 0)]^3$ . Note that both  $u(x, t)$  and  $h(x, t)$  do not depend on  $\alpha$ .

These functions also appear in the analysis of the statistics of time integrated current in the symmetric exclusion process on an infinite line (see Appendix D).

In order to compute the fourth cumulant we must find  $R_4$  and  $Y_3$ . Combining the series expansion (106)–(107) and (113) we get

$$R_4 = \int_{-\infty}^{\infty} dx U_4(x) + \int_0^T dt \int_{-\infty}^{\infty} dx V_4(x, t)$$

with

$$\begin{aligned}
U_4 &= \frac{[q_2(x, 0)]^2 + 2q_1(x, 0)q_3(x, 0)}{2\rho(1 - \alpha\rho)} - \frac{(1 - 2\alpha\rho)}{2\rho^2(1 - \alpha\rho)^2} (q_1(x, 0))^2 q_2(x, 0) \\
&\quad + \frac{(1 - 2\alpha\rho)^2}{12\rho^3(1 - \alpha\rho)^3} (q_1(x, 0))^4 + \frac{\alpha}{12\rho^2(1 - \alpha\rho)^2} (q_1(x, 0))^4 \\
V_4 &= \rho(1 - \alpha\rho) \left[ (\partial_x p_2)^2 + 2(\partial_x p_1)(\partial_x p_3) \right] - \alpha q_1^2 (\partial_x p_1)^2 \\
&\quad + (1 - 2\alpha\rho) \left[ q_2(\partial_x p_1)^2 + 2q_1(\partial_x p_1)(\partial_x p_2) \right].
\end{aligned}$$

We now again express  $R_4$  through the corresponding solution for  $\alpha = 0$  case:

$$R_4 = (1 - \alpha\rho)(1 - 2\alpha\rho)^2 \left[ \widehat{R}_4 \Big|_{\widehat{Y}_1 \rightarrow \frac{Y_1}{1-2\alpha\rho}} \right] + \rho^2 \alpha (1 - \alpha\rho)^2 \mathcal{I}. \quad (127)$$

The last term is given by

$$\begin{aligned}
\mathcal{I} &= \int_{-\infty}^{\infty} dx \left[ \left( \frac{\widehat{q}_1(x, 0)}{\rho} \right) \left( \frac{h(x, 0)}{\rho^2} \right) + \frac{1}{12} \left( \frac{\widehat{q}_1(x, 0)}{\rho} \right)^4 \right] \\
&\quad + \int_0^T dt \int_{-\infty}^{\infty} dx \left[ 2(\partial_x \widehat{p}_1) \left( \frac{\partial_x u}{\rho} \right) - \left( \frac{\widehat{q}_1}{\rho} \right)^2 (\partial_x \widehat{p}_1)^2 \right].
\end{aligned}$$

This only involves the solution for the  $\alpha = 0$  case.

The second quantity required to compute the fourth cumulant is  $Y_3$ . Combining the series expansion (99) and (51), and using the relation of the optimal fields to their counterparts for  $\alpha = 0$ , we can express  $Y_3$  in terms of the hat variables:

$$Y_3 = (1 - \alpha\rho)(1 - 2\alpha\rho)^2 \widehat{Y}_3 \Big|_{\widehat{Y}_1 \rightarrow \frac{Y_1}{1-2\alpha\rho}} + \alpha\rho(1 - \alpha\rho)^2 \int_0^{\infty} dx \left[ \frac{h(x, T)}{\rho^2} - \frac{h(x, 0)}{\rho^2} \right]. \quad (128)$$

Combining this equation with (127), we find that  $Y_3 - R_4/\rho$  which appears in the fourth cumulant (103) can be written as

$$\begin{aligned}
-\frac{R_4}{\rho} + Y_3 &= (1 - \alpha\rho)(1 - 2\alpha\rho)^2 \left[ -\frac{\widehat{R}_4}{\rho} + \widehat{Y}_3 \right] \Big|_{\widehat{Y}_1 \rightarrow \frac{Y_1}{1-2\alpha\rho}} \\
&\quad + \alpha\rho(1 - \alpha\rho)^2 \left[ -\mathcal{I} + \int_0^{\infty} dx \left( \frac{h(x, T)}{\rho^2} - \frac{h(x, 0)}{\rho^2} \right) \right].
\end{aligned} \quad (129)$$

The term inside the square brackets in the bottom line of (129) simplifies to

$$-\mathcal{I} + \int_0^{\infty} dx \left( \frac{h(x, T)}{\rho^2} - \frac{h(x, 0)}{\rho^2} \right) = \frac{1}{4!} \left( \frac{8 - 6\sqrt{2}}{\sqrt{\pi}} \right) \sqrt{T}. \quad (130)$$

This identity was verified numerically using Mathematica. We also obtained an analytical proof by comparing to the analysis of the current in the SEP with uniform initial profile (see Appendix D).

On the other hand, from the exact solution (54) for the  $\alpha = 0$  case we get (by substituting  $Y(B)$  from (99) in (54) and expanding in powers of  $B$ )

$$-\frac{\widehat{R}_4}{\rho} + \widehat{Y}_3 = \frac{\sqrt{T}}{12\sqrt{\pi}} - \frac{\widehat{Y}_1}{6} + \frac{\widehat{Y}_1^2}{4\sqrt{T}\pi}. \quad (131)$$

This allows us to simplify the term inside the square brackets in the top line of (129)

$$\left[ -\frac{\widehat{R}_4}{\rho} + \widehat{Y}_3 \right] \Big|_{\widehat{Y}_1 \rightarrow \frac{Y_1}{1-2\alpha\rho}} = \frac{\sqrt{T}}{12\sqrt{\pi}} - \frac{Y_1}{6(1-2\alpha\rho)} + \frac{Y_1^2}{4\sqrt{T}\pi(1-2\alpha\rho)^2}. \quad (132)$$

Substituting  $Y_1$  from (120), we obtain

$$\frac{1}{\sqrt{T}} \left[ -\frac{\widehat{R}_4}{\rho} + \widehat{Y}_3 \right] \Big|_{\widehat{Y}_1 \rightarrow \frac{Y_1}{1-2\alpha\rho}} = \frac{1}{12\sqrt{\pi}} - \frac{(1-\alpha\rho)}{(1-2\alpha\rho)} \frac{1}{3\sqrt{\pi}} + \frac{(1-\alpha\rho)^2}{(1-2\alpha\rho)^2} \frac{1}{\pi^{3/2}}. \quad (133)$$

After all this work Eq. (129) becomes

$$\begin{aligned} \frac{1}{\sqrt{T}} \left[ -\frac{R_4}{\rho} + Y_3 \right] &= (1-\alpha\rho)(1-2\alpha\rho)^2 \frac{1}{12\sqrt{\pi}} - (1-\alpha\rho)^2(1-2\alpha\rho) \frac{1}{3\sqrt{\pi}} \\ &\quad + (1-\alpha\rho)^3 \frac{1}{\pi^{3/2}} + \frac{\alpha\rho(1-\alpha\rho)^2}{4!} \left( \frac{8-6\sqrt{2}}{\sqrt{\pi}} \right). \end{aligned} \quad (134)$$

Plugging (134) and  $\lambda_1 = \rho$  into (103) we obtain

$$\langle X_T^4 \rangle_c = \frac{(1-\alpha\rho)}{\rho^3} \left[ 1 - \left( 4 - (8-3\sqrt{2})\alpha\rho \right) (1-\alpha\rho) + \frac{12}{\pi} (1-\alpha\rho)^2 \right] \frac{2\sqrt{T}}{\sqrt{\pi}}. \quad (135)$$

Specifying (135) to  $\alpha = 0$  we get back the result (66), whereas setting  $\alpha = 1$  we arrive at the fourth cumulant for the SEP

$$\langle X_T^4 \rangle_c = \frac{1-\rho}{\rho^3} \left[ 1 - \left( 4 - (8-3\sqrt{2})\rho \right) (1-\rho) + \frac{12}{\pi} (1-\rho)^2 \right] \frac{2\sqrt{T}}{\sqrt{\pi}}. \quad (136)$$

This result, announced in Ref. [29], is valid in the annealed setting.<sup>7</sup> When  $\rho \rightarrow 1$ , this expression matches the result derived in Ref. [41].

*Remark:* The fourth cumulant in the quenched setting can be calculated along similar lines to yield

$$\langle X_T^4 \rangle_c^{\mathcal{Q}} = \frac{1-\alpha\rho}{\rho^3} \left[ 2(1-2\alpha\rho)^2 \left( \frac{9}{\pi} \arctan \left( \frac{1}{2\sqrt{2}} \right) - 1 \right) + \alpha\rho(1-\alpha\rho)(4-3\sqrt{2}) \right] \sqrt{\frac{2}{\pi}} \sqrt{T}.$$

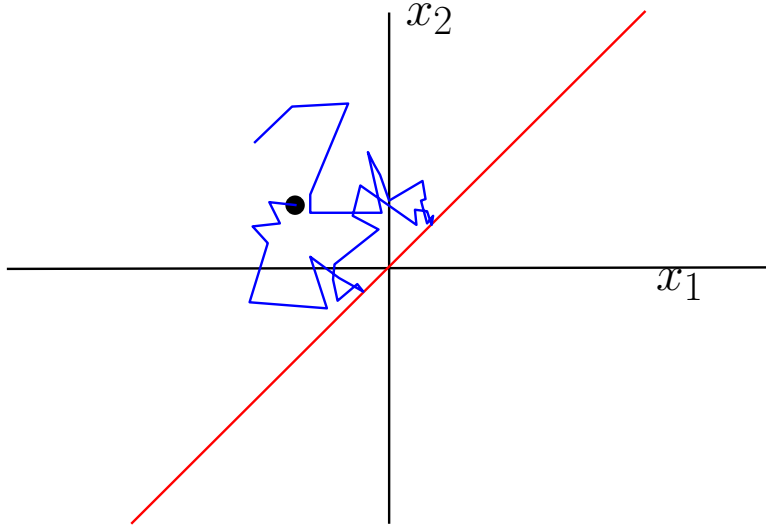
For Brownian particles,  $\alpha = 0$ , we recover Eq. (64). Setting  $\alpha = 1$ , we obtain the fourth cumulant for the SEP in the quenched case:

$$\langle X_T^4 \rangle_c^{\mathcal{Q}} = (1-\rho) \left[ 2 \frac{(1-2\rho)^2}{\rho^3} \left( \frac{9}{\pi} \arctan \left( \frac{1}{2\sqrt{2}} \right) - 1 \right) + \frac{1-\rho}{\rho^2} (4-3\sqrt{2}) \right] \sqrt{\frac{2}{\pi}} \sqrt{T}.$$

## 6 A microscopic derivation for the Brownian point particles.

The microscopic problem was first studied by Harris [17], who derived an exact formula for the variance of the tagged particle. The analysis used the fact that trajectories of the particles are related to the trajectories of non-interacting particles with an exchange of particle index to keep the ordering same (see [77] for a recent reference).

There is an equivalent description of the problem in terms of the phase space trajectories [20,25,80]. Consider  $2n+1$  point particles diffusing on a one-dimensional line. The only interaction between particles is the hard-core repulsion which preserves the order of the particles. The particles are indexed by  $\{-n, -n+1, \dots, n\}$ . The central particle is set to be the tagged particle. Let  $\mathbf{Y} \equiv \{y_{-n}, \dots, y_n\}$ ,



**Fig. 6** A schematic representation of a sample trajectory of two one-dimensional Brownian point particles on the coordinate plane  $(x_1, x_2)$ . Due to the hard-core repulsion between particles the motion is confined in the domain  $x_2 > x_1$ , whose boundary is denoted by the diagonal line.

with  $-L \leq y_{-n} < \dots < y_n \leq L$ , be the positions at time  $t = 0$ . The positions at time  $t$  are denoted by  $\mathbf{X} \equiv \{x_{-n}, \dots, x_n\}$ .

The evolution of the particles can be described by diffusion in a  $2n + 1$  dimensional space confined to the chamber  $x_{-n} < x_{-n+1} < \dots < x_n$  (see Figure 6 for a schematic). The probability  $P_t(\mathbf{X}|\mathbf{Y})$  of the particle position follows a diffusion equation

$$\partial_t P_t(\mathbf{X}|\mathbf{Y}) = \sum_{i=-n}^n \partial_{x_i}^2 P_t(\mathbf{X}|\mathbf{Y}) \quad (137)$$

where the diffusion coefficient is again set to unity. The single-file constraint is implemented by a reflecting boundary condition along the boundary of the chamber:

$$\partial_{x_i} P_t(\mathbf{X}|\mathbf{Y}) = \partial_{x_{i+1}} P_t(\mathbf{X}|\mathbf{Y}) \quad \text{at } x_i = x_{i+1} \text{ for all } -n \leq i < n. \quad (138)$$

The solution to Eqs. (137)–(138) can be written as

$$P_t(\mathbf{X}|\mathbf{Y}) = \int_{-\infty}^{\infty} \frac{dk_{-n}}{2\pi} \dots \int_{-\infty}^{\infty} \frac{dk_n}{2\pi} \sum_{\sigma} \prod_{j=-n}^n e^{-tk_j^2 + ik_j(x_{\sigma(j)} - y_j)}, \quad (139)$$

where  $\sigma$  is the permutation operator acting on the indexes of the particles. The above solution is applicable only within the chamber  $x_{-n} < x_{-n+1} < \dots < x_n$ . Equation (139) is one of the simplest examples of the Bethe ansatz.

The Gaussian integrals in (139) can be evaluated yielding

$$P_t(\mathbf{X}|\mathbf{Y}) = \sum_{\sigma} \prod_{j=-n}^n g_t(x_j | y_{\sigma(j)}), \quad g_t(x|y) \equiv \frac{1}{\sqrt{4\pi t}} e^{-\frac{(x-y)^2}{4t}}. \quad (140)$$

---

<sup>7</sup> We also derived (136) by a straightforward perturbative calculation, without introducing the interpolating parameter  $\alpha$ .



### 6.1 Probability of the tagged particle position

The central particle ( $j = 0$ ) is chosen to be the tagged particle. Without loss of generality, we assume that the tagged particle starts at the origin,  $y_0 = 0$ . The probability of finding the tagged particle at  $x_0 = x$  at time  $t$  is

$$\text{Prob}_t(x|\mathbf{Y}) = \int_{-\infty}^{x-n+1} dx_{-n} \int_{-\infty}^{x-n+2} dx_{-n+1} \dots \int_{-\infty}^x dx_{-1} \int_x^\infty dx_1 \int_{x_1}^\infty dx_2 \dots \int_{x_{n-1}}^\infty dx_n P_t(\mathbf{X}|\mathbf{Y}).$$

This function has been studied in great details by Rödénbeck *et al.* [20] for the annealed initial condition. Their results allow one to extract the large deviation function. Here we present a short alternative derivation, and then we determine the large deviation in the quenched case.

Let

$$G(x, y_1, \dots, y_n) = \int_x^\infty dz_1 \int_{z_1}^\infty dz_2 \dots \int_{z_{n-1}}^\infty dz_n \prod_{j=1}^n g_t(z_j|y_j). \quad (141)$$

In terms of this function

$$\text{Prob}_t(x|\mathbf{Y}) = \sum_{\sigma} g_t(x|y_{\sigma(0)}) G(-x, -y_{\sigma(-1)}, -y_{\sigma(-2)}, \dots, -y_{\sigma(-n)}) G(x, y_{\sigma(1)}, y_{\sigma(2)}, \dots, y_{\sigma(n)}).$$

Let us rewrite the above formula as

$$\text{Prob}_t(x|\mathbf{Y}) = \sum_{k=-n}^n g_t(x|y_k) A_k[x, \mathbf{Y}_k], \quad (142)$$

where we grouped the terms with same values of  $\sigma(0) = k$ . Further,  $\mathbf{Y}_k$  denotes the subset of  $\mathbf{Y}$  excluding the  $k$ -th element  $y_k$ . The amplitude is

$$A_k[x, \mathbf{Y}_k] = \sum_{\sigma_k} G(-x, -y_{\sigma_k(-1)}, -y_{\sigma_k(-2)}, \dots, -y_{\sigma_k(-n)}) G(x, y_{\sigma_k(1)}, y_{\sigma_k(2)}, \dots, y_{\sigma_k(n)}), \quad (143)$$

where the permutation operator  $\sigma_k$  denotes permutations acting on the subset  $\mathbf{Y}_k$ .

Equation (143) can be simplified using an identity (174), proved in the Appendix E, resulting in

$$A_k[x, \mathbf{Y}_k] = \sum_{\epsilon_{-n}} \dots \sum_{\epsilon_{k-1}} \sum_{\epsilon_{k+1}} \dots \sum_{\epsilon_n} \left[ \delta_{\sum_{j \neq k} \epsilon_j, 0} \right] \prod_{\ell \neq k} \frac{1}{2} \text{erfc} \left( \epsilon_\ell \frac{x - y_\ell}{\sqrt{4t}} \right), \quad (144)$$

where  $\epsilon_\ell$  are binary variables taking values  $\pm 1$ .

### 6.2 Comparison with a random field Ising model

The appearance of binary variables  $\epsilon_j$  in Eq. (144) suggests to seek a connection to an Ising model with non-interacting spins  $\epsilon_j$ . Such a connection indeed exists. To see it we use an identity

$$\text{erfc} \left( \epsilon_j \frac{x - y_j}{\sqrt{4t}} \right) = \frac{\exp(-h_j \epsilon_j)}{\cosh(h_j)}, \quad (145)$$

where  $h_j$  is the magnetic field acting on the spin  $\epsilon_j$ , defined as

$$h_j = \text{arctanh} \left[ \text{erf} \left( \frac{x - y_j}{\sqrt{4t}} \right) \right]. \quad (146)$$

This leads to an expression for the amplitude

$$A_k[x, \mathbf{Y}_k] = \frac{Z[\mathbf{h}_k]}{\prod_{j \neq k} 2 \cosh(h_j)}, \quad (147)$$

where  $Z[\mathbf{h}_k]$  is the partition function of an Ising chain of  $2n$  spins in the ensemble with total magnetization zero, with  $\mathbf{h}_k = (h_{-n}, \dots, h_{k-1}, h_{k+1}, \dots, h_n)$ . The partition function is defined as

$$Z[\mathbf{h}_k] = \sum_{\epsilon_{-n}} \cdots \sum_{\epsilon_{k-1}} \sum_{\epsilon_{k+1}} \cdots \sum_{\epsilon_n} \left[ \delta_{\sum_{j \neq k} \epsilon_j, 0} \right] \prod_{\ell \neq k} e^{h_\ell \epsilon_\ell}. \quad (148)$$

The spins do not interact with each other. In terms of the partition function (148) the probability of tagged particle position can be written as

$$\text{Prob}_t(x|\mathbf{Y}) = \sum_{k=-n}^n \frac{g_t(x|y_k)}{\prod_{j \neq k} 2 \cosh(h_j)} Z[\mathbf{h}_k]. \quad (149)$$

In this formulation in terms of the Ising model, the quenched case where the initial particle positions  $y_j$  are fixed corresponds to a quenched magnetic field. The annealed case where the  $y_j$  are fluctuating corresponds to the fluctuating magnetic field. It is well known that the properties of a random field Ising model is different in the quenched and in the annealed ensemble. Then, the difference in the statistics of the tagged particle position between the quenched and the annealed initial condition can be related to the non-equivalence of the random field Ising model in the two ensembles.

### 6.3 Normal diffusion

There is a finite number of particles,  $2n + 1$  in our setting, diffusing on an infinite line. Therefore the tagged particle has a normal diffusion in the large time limit. This large time is set by  $t \gg L^2$ , since initially particles are distributed within an interval  $[-L, L]$ .

In this limit, the leading behavior of the probability given by (149) comes from replacing  $y_j = 0$  for all  $j$ . This corresponds to the case where the magnetic field  $h_j \equiv h = \text{arctanh}[\text{erf}(x/\sqrt{4t})]$  is uniform. The corresponding partition function is  $Z[\mathbf{h}_k] = \frac{(2n)!}{n!n!}$ . Note that it is independent of  $h$  because the total magnetization vanishes.

Substituting  $y_j = 0$  for all  $j$  into (149) we obtain

$$\text{Prob}_t(x|\mathbf{Y}) \simeq g_t(x|0) \left[ \frac{1}{4} \text{erfc}\left(\frac{x}{\sqrt{4t}}\right) \text{erfc}\left(-\frac{x}{\sqrt{4t}}\right) \right]^n \frac{(2n+1)!}{n!n!}. \quad (150)$$

The result does not depend on the initial positions  $\mathbf{Y}$  of the particles, that is whether the initial state is annealed or quenched. For large  $n$  the leading dependence on  $n$  becomes

$$\text{Prob}_t(x = \xi\sqrt{4t}) \simeq \frac{\sqrt{n}}{\pi\sqrt{t}} e^{n\psi(\xi)}, \quad (151)$$

with  $\psi(\xi) = \ln[\text{erfc}(\xi)\text{erfc}(-\xi)]$ . At large times, the distribution is expected to be Gaussian. This can be confirmed by expanding  $\psi(\xi)$  for small  $\xi$  to give

$$\text{Prob}_t(x) \simeq \frac{1}{\sqrt{4\pi\mathcal{D}t}} e^{-\frac{x^2}{4\mathcal{D}t}}, \quad (152)$$

with the self-diffusion constant  $\mathcal{D} = \frac{\pi}{4n}$ . The self-diffusion constant decreases as  $n$  increases indicating the sub-diffusive behavior observed in the macroscopic calculation.

### 6.4 Sub-diffusion

This limit is defined by the number of particles  $n \rightarrow \infty$  and  $L \rightarrow \infty$  keeping the density  $n/L = \rho$  constant. The central (tagged) particle is caged by infinitely many particles. From the hydrodynamic result in earlier sections it is expected that in this limit the probability of the tagged particle position has the large deviation form  $\text{Prob}_t(x) \asymp \exp[-\sqrt{4t} \phi(x/\sqrt{4t})]$ . We now derive  $\phi(\xi)$  for quenched and annealed settings, and compare with the results from the hydrodynamic approach.

### Quenched Case

As a quenched initial state we choose the equidistant one: All particles at  $t = 0$  are placed deterministically with separation  $\rho^{-1}$  between adjacent particles. Then the initial position of the  $\ell^{\text{th}}$  particle is  $y_\ell = \rho^{-1}\ell$ . We write the probability given by Eq. (142) as

$$\text{Prob}_t(x) = P_0(x) + P_-(x) + P_+(x), \quad (153)$$

with

$$P_0(x) = g_t(x|0)A_0[x, \mathbf{Y}_0], \quad (154)$$

$$P_-(x) = \sum_{k=-1}^{-n} g_t(x|y_k)A_k[x, \mathbf{Y}_k], \quad (155)$$

$$P_+(x) = \sum_{k=1}^n g_t(x|y_k)A_k[x, \mathbf{Y}_k]. \quad (156)$$

All the three terms  $P_0$ ,  $P_\pm$  have the same asymptotic large deviation form at large  $t$  limit. They only differ in the sub-leading terms in  $t$ . The advantage of using the representation (153) is that it is much simpler to analyze the term  $P_0$  and to extract the large deviation function. We derive the large deviation function using a saddle point analysis of  $P_0$ .

We use the expression for  $A_0$  from (144) which leads to

$$P_0(x) = g_t(x|0) \sum_{\epsilon_{-n}} \cdots \sum_{\epsilon_{-1}} \sum_{\epsilon_1} \cdots \sum_{\epsilon_n} \left[ \delta_{\sum_{j \neq 0} \epsilon_j, 0} \right] \prod_{\ell \neq 0} \frac{1}{2} \text{erfc} \left( \epsilon_\ell \frac{x - \rho^{-1}\ell}{\sqrt{4t}} \right). \quad (157)$$

We have used the quenched initial particle position  $y_\ell = \rho^{-1}\ell$ . The Kronecker delta function can be replaced by an integral,

$$P_0(x) = g_t(x|0) \int dB \sum_{\epsilon_{-n}} \cdots \sum_{\epsilon_{-1}} \sum_{\epsilon_1} \cdots \sum_{\epsilon_n} \prod_{\ell \neq 0} \left[ e^{\epsilon_\ell B/2} \right] \frac{1}{2} \text{erfc} \left( \epsilon_\ell \frac{x - \rho^{-1}\ell}{\sqrt{4t}} \right),$$

where  $B$  is the integration variable and the factor  $1/2$  is for later convenience. In this form, the binary variables  $\epsilon_\ell$  are decoupled and their sums can be performed rather easily, leading to

$$P_0(x) = g_t(x|0) \int dB \exp \left[ \sum_{\ell=1}^n \log \left\{ 1 + (e^B - 1) \frac{1}{2} \text{erfc} \left( \frac{x + \rho^{-1}\ell}{\sqrt{4t}} \right) \right\} \right. \\ \left. + \sum_{\ell=1}^n \log \left\{ 1 + (e^{-B} - 1) \frac{1}{2} \text{erfc} \left( \frac{\rho^{-1}\ell - x}{\sqrt{4t}} \right) \right\} \right].$$

Taking the limit  $n \rightarrow \infty$  with uniform density of particles  $\rho$  and replacing the summation over  $\ell$  by an integral, yields,

$$P_0(x) = g_t(x|0) \int dB \exp \left[ \rho \int_0^\infty dz \log \left\{ 1 + (e^B - 1) \frac{1}{2} \text{erfc} \left( \frac{x + z}{\sqrt{4t}} \right) \right\} \right. \\ \left. + \rho \int_0^\infty dz \log \left\{ 1 + (e^{-B} - 1) \frac{1}{2} \text{erfc} \left( \frac{z - x}{\sqrt{4t}} \right) \right\} \right].$$

To compute the large deviation function we write  $\xi = x/\sqrt{4t}$  and recast the above formula into

$$P_0(\xi\sqrt{4t}) = \frac{\exp(-\xi^2)}{\sqrt{4\pi t}} \int dB \exp \left[ \rho\sqrt{4t} \int_\xi^\infty dz \log \left\{ 1 + (e^B - 1) \frac{1}{2} \text{erfc}(z) \right\} \right. \\ \left. + \rho\sqrt{4t} \int_{-\xi}^\infty dz \log \left\{ 1 + (e^{-B} - 1) \frac{1}{2} \text{erfc}(z) \right\} \right].$$

At large  $t$  the integral over  $B$  is dominated by the saddle point leading to a parametric solution of the large deviation function

$$\begin{aligned} \phi_{\mathcal{Q}}(\xi) = - \lim_{t \rightarrow \infty} \frac{\log P_0(\xi \sqrt{4t})}{\sqrt{4t}} &= -\rho \int_{\xi}^{\infty} dz \log \left\{ 1 + (e^B - 1) \frac{1}{2} \operatorname{erfc}(z) \right\} \\ &\quad - \rho \int_{-\xi}^{\infty} dz \log \left\{ 1 + (e^{-B} - 1) \frac{1}{2} \operatorname{erfc}(z) \right\} \end{aligned} \quad (158)$$

with  $B$  determined from the optimization requirement  $\frac{d\phi_{\mathcal{Q}}(\xi)}{dB} = 0$ . These results are identical to (47)–(48) obtained using the MFT.

### Annealed Case

In this case, the initial positions of the particles are uniformly distributed in the interval  $[-L, L]$ . Averaging over the initial positions the probability of the tagged particle position can be expressed as

$$\begin{aligned} \operatorname{Prob}_t(x) = & \left[ \frac{n!}{L^n} \int_{-L}^0 dy_{-1} \int_{-L}^{y_{-1}} dy_{-2} \cdots \int_{-L}^{y_{-n+1}} dy_{-n} \right] \\ & \left[ \frac{n!}{L^n} \int_0^L dy_1 \int_{y_1}^L dy_2 \cdots \int_{y_{n-1}}^L dy_n \right] \operatorname{Prob}_t(x|\mathbf{Y}), \end{aligned} \quad (159)$$

where  $\operatorname{Prob}_t(x|\mathbf{Y})$  is given in (142). As in the quenched case, we write  $\operatorname{Prob}_t(x)$  as a sum of three terms. All terms have the same large deviation form. Hence it is sufficient to analyze

$$\begin{aligned} \operatorname{Prob}_t(x) \sim P_0(x) = g_t(x|0) & \left[ \frac{n!}{L^n} \int_{-L}^0 dy_{-1} \int_{-L}^{y_{-1}} dy_{-2} \cdots \int_{-L}^{y_{-n+1}} dy_{-n} \right] \\ & \times \left[ \frac{n!}{L^n} \int_0^L dy_1 \int_{y_1}^L dy_2 \cdots \int_{y_{n-1}}^L dy_n \right] \\ & \times \sum_{\epsilon_{-n}} \cdots \sum_{\epsilon_{-1}} \sum_{\epsilon_1} \cdots \sum_{\epsilon_n} \left[ \delta_{\sum_{j \neq 0} \epsilon_j, 0} \right] \prod_{\ell \neq 0} \frac{1}{2} \operatorname{erfc} \left( \epsilon_{\ell} \frac{x - y_{\ell}}{\sqrt{4t}} \right). \end{aligned}$$

The symmetry of the integrand allows us to re-write it as

$$\begin{aligned} P_0(x) = g_t(x|0) & \sum_{\epsilon_{-n}} \cdots \sum_{\epsilon_{-1}} \sum_{\epsilon_1} \cdots \sum_{\epsilon_n} \left[ \delta_{\sum_{j \neq 0} \epsilon_j, 0} \right] \\ & \left[ \prod_{\ell=-1}^{-n} \frac{1}{L} \int_{-L}^0 dy_{\ell} \frac{1}{2} \operatorname{erfc} \left( \epsilon_{\ell} \frac{x - y_{\ell}}{\sqrt{4t}} \right) \right] \left[ \prod_{\ell=1}^n \frac{1}{L} \int_0^L dy_{\ell} \frac{1}{2} \operatorname{erfc} \left( \epsilon_{\ell} \frac{x - y_{\ell}}{\sqrt{4t}} \right) \right]. \end{aligned}$$

The large deviation function can be derived from the above expression using a saddle point approximation. Replacing the delta function by an integral representation we get

$$\begin{aligned} P_0(x) = g_t(x|0) \int dB & \left[ \prod_{\ell=-1}^{-n} \sum_{\epsilon_{\ell}} e^{\epsilon_{\ell} B/2} \frac{1}{L} \int_{-L}^0 dy_{\ell} \frac{1}{2} \operatorname{erfc} \left( \epsilon_{\ell} \frac{x - y_{\ell}}{\sqrt{4t}} \right) \right] \\ & \left[ \prod_{\ell=1}^n \sum_{\epsilon_{\ell}} e^{\epsilon_{\ell} B/2} \frac{1}{L} \int_0^L dy_{\ell} \frac{1}{2} \operatorname{erfc} \left( \epsilon_{\ell} \frac{x - y_{\ell}}{\sqrt{4t}} \right) \right]. \end{aligned}$$

Completing the summation over the binary variables  $\epsilon_{\ell}$  we get

$$\begin{aligned} P_0(x) = g_t(x|0) \int dB & \left[ 1 + (e^B - 1) \frac{1}{L} \int_{-L}^0 dy \frac{1}{2} \operatorname{erfc} \left( \frac{x - y}{\sqrt{4t}} \right) \right]^n \\ & \left[ 1 + (e^{-B} - 1) \frac{1}{L} \int_0^L dy \frac{1}{2} \operatorname{erfc} \left( \frac{y - x}{\sqrt{4t}} \right) \right]^n, \end{aligned}$$

where we have replaced  $y_\ell$  by  $y$ , as the former reduces to a dummy variable after the summation.

In the limit of  $n \rightarrow \infty$  and  $L \rightarrow \infty$  keeping  $n/L = \rho$  fixed, the above expression results in

$$P_0(x) = g_t(x|0) \int dB \exp \left[ \rho (e^B - 1) \int_{-\infty}^0 dy \frac{1}{2} \operatorname{erfc} \left( \frac{x-y}{\sqrt{4t}} \right) + \rho (e^{-B} - 1) \int_0^{\infty} dy \frac{1}{2} \operatorname{erfc} \left( \frac{y-x}{\sqrt{4t}} \right) \right].$$

In terms of rescaled coordinates  $\xi = x/\sqrt{4t}$  and  $z = y/\sqrt{4t}$  the above can be written as

$$P_0(\xi\sqrt{4t}) = \frac{e^{-\xi^2}}{\sqrt{4\pi t}} \int dB \exp \left[ \rho \sqrt{t} \left\{ (e^B - 1) \int_{-\infty}^0 dz \operatorname{erfc}(\xi - z) + (e^{-B} - 1) \int_0^{\infty} dz \operatorname{erfc}(z - \xi) \right\} \right].$$

At large  $t$ , the integral is dominated by the saddle point. This leads to the large deviation function

$$\phi_{\mathcal{A}}(\xi) \simeq - \lim_{t \rightarrow \infty} \frac{\log P_0(\xi\sqrt{4t})}{\sqrt{4t}} = -\rho \frac{(e^B - 1)}{2} \int_{\xi}^{\infty} dz \operatorname{erfc}(z) - \rho \frac{(e^{-B} - 1)}{2} \int_{-\infty}^{\xi} dz \operatorname{erfc}(z), \quad (160)$$

with  $B$  determined from the saddle point condition

$$\frac{d\phi_{\mathcal{A}}(\xi)}{dB} = 0. \quad (161)$$

To show the equivalence with the MFT result (59) we use (161) to deduce

$$e^{2B} = \frac{\int_{-\xi}^{\infty} dz \operatorname{erfc}(z)}{\int_{\xi}^{\infty} dz \operatorname{erfc}(z)}. \quad (162)$$

Combining this with (160) one recovers the large deviation function (59).

## 7 Summary

We studied the full statistics of the displacement of the tagged particle in single-file diffusion. Our analysis mostly relies on the macroscopic fluctuation theory (MFT). We found the full solution for the simplest single-file system composed of impenetrable Brownian particles (we computed the optimal paths, the large deviation function, etc.). This single-file system is also amenable to an exact microscopic analysis. In the annealed case, the large deviation function was originally computed in [20] and we presented another (shorter) derivation; the quenched case has not been studied before. The predictions based on the exact analyses and those which were derived using the MFT match for the single-file system composed of Brownian particles.

Single-file systems with arbitrary transport coefficients cannot be solved exactly. Yet such general systems are tractable perturbatively. Specifically, we performed an expansion in the powers of the fugacity and derived an exact formula for the variance of position of the tagged particle valid in the general case of arbitrary  $D(\rho)$  and  $\sigma(\rho)$ . The forth cumulant can also be computed for a fairly general class of single-file systems, e.g., for systems with constant diffusion coefficient and quadratic (in density) mobility. The well-known example of such system is the SEP, and we computed the forth cumulant of the tagged particle position for the SEP.

Our study can be extended in many directions. One should be able to use the MFT to calculate statistical properties of the tagged particle trajectory (the two-time correlation functions [69], the first passage times etc.). It would also be interesting to investigate the influence of a bias in the simple exclusion process. If one considers an asymmetrically hopping tagged particle in a bath of SEP, all the cumulants scale as  $\sqrt{T}$  [81, 82]. The full statistics has been recently determined in the high density limit [41] when the leading asymptotic can be extracted by treating the vacancies as independent random walkers; for the finite density even the variance remains unknown. If all particles are biased, even the scale of fluctuations depend on the setting: In the annealed case, the position fluctuation of the tagged particle grows as  $\sqrt{T}$  [9, 13]; in the quenched case, the exponent changes from 1/2 to 1/3 [8, 11, 14, 83–85]. It is not known whether a hydrodynamic approach based on the MFT can

capture these behaviors and if the Tracy-Widom distribution, which is ubiquitous in these problems (see e.g. [22]), can be retrieved as a solution of some optimal path equations.

From a more general point of view, it would be interesting to give a physical interpretation of the conjugate field that appears in the MFT optimal path equations and to classify the one-dimensional gases that could lead to classically integrable partial differential equations, that can be solved using the inverse scattering method [86].

**Acknowledgements** We benefitted from discussions with A. Dhar, S. Majumdar, S. Mallick, B. Meerson, S. Sabhapandit and R. Voituriez. We are also grateful to the referees for very useful comments and suggestions. This research was partly supported by grant No. 2012145 from the BSF. We thank the Galileo Galilei Institute for Theoretical Physics for excellent working conditions and the INFN for partial support.

## A Functional derivative of $Y[q]$

The tagged particle position is a functional  $Y[q]$  of the initial  $q(x, 0)$  and the final  $q(x, T)$  density profile, defined by the relation

$$\int_0^{Y[q]} dx q(x, T) = \int_0^\infty dx \left[ q(x, T) - q(x, 0) \right]. \quad (163)$$

which comes from equation (6). This way of writing makes both the integrals convergent.

The functional derivatives appearing in the boundary conditions (15)–(17) can be easily computed from the definition above. To proceed, consider a small variation in the final density  $q(x, T) \rightarrow q(x, T) + \delta q(x, T)$  leading to a change  $Y \rightarrow Y + \delta Y$ . Corresponding to this variation the above equation becomes

$$\int_0^{Y+\delta Y} dx \left[ q(x, T) + \delta q(x, T) \right] = \int_0^\infty dx \left[ q(x, T) - q(x, 0) \right] + \int_0^\infty dx \delta q(x, T).$$

Using the formula (163) this reduces to,

$$\int_Y^{Y+\delta Y} dx q(x, T) = \int_{Y+\delta Y}^\infty dx \delta q(x, T). \quad (164)$$

Assuming that the change  $\delta Y$  is small for a small variation  $\delta q(x, T)$  and keeping only the linear terms, the above leads to

$$\delta Y = \int_{-\infty}^\infty dx \left[ \frac{\Theta(x - Y)}{q(Y, T)} \right] \delta q(x, T). \quad (165)$$

Hence, we have

$$\frac{\delta Y}{\delta q(x, T)} = \frac{\Theta(x - Y)}{q(Y, T)}. \quad (166)$$

The functional derivative with respect to the initial density  $q(x, 0)$  is similarly derived:

$$\frac{\delta Y}{\delta q(x, 0)} = -\frac{\Theta(x)}{q(Y, T)}. \quad (167)$$

## B Inequality between cumulants in the annealed and quenched cases

Because of the fluctuations in the initial state, the tagged particle is expected to have larger displacements in the annealed case than in the quenched case. More generally for any single file system

$$\mu_{\mathcal{A}}(\lambda) \geq \mu_{\mathcal{Q}}(\lambda) \quad (168)$$

which implies  $\langle X_T^2 \rangle_{\mathcal{A}} \geq \langle X_T^2 \rangle_{\mathcal{Q}}$ .

To derive (168) we recall that the two initial settings differ by how the average over initial state is taken:

$$\mu_{\mathcal{A}}(\lambda) = \log \left\langle e^{\lambda X_T} \right\rangle_{\text{evolution} + \text{initial}}, \quad (169)$$

$$\mu_{\mathcal{Q}}(\lambda) = \left\langle \log \left\langle e^{\lambda X_T} \right\rangle_{\text{evolution}} \right\rangle_{\text{initial}}. \quad (170)$$

Here the subscript *evolution* denotes average over stochastic evolution of the system and the subscript *initial* denotes average over initial state. Equation (168) then follows from the Jensen inequality [87] because log is a concave function. The inequality (168) implies the opposite relation for the large deviation functions and therefore explains why the large deviation function in the quenched case exceeds the large deviation function in the annealed case (Fig. 4).

## C Mobility $\sigma(\rho)$

We first derive  $\sigma(\rho)$  for the Brownian particles with hard core repulsion. Let there are  $n$  number of particles at equilibrium within an interval of length  $L$ . All particle positions within the interval are equally probable. This leads to the canonical free energy density

$$f\left(\frac{n}{L}\right) = -L^{-1} \log\left(\frac{L^n}{n!}\right).$$

In the limit  $n \rightarrow \infty$  and  $L \rightarrow \infty$  with finite  $n/L = \rho$ , the above formula becomes  $f(\rho) = \rho \ln \rho - \rho$ . Using it together with the fluctuation-dissipation relation (5) one gets  $\sigma(\rho) = \frac{2D(\rho)}{f''(\rho)} = 2\rho$ .

Consider now the SEP with  $n$  particles on the ring of  $L$  sites. In the equilibrium, all configurations are equally probable which leads to the free energy density

$$f\left(\frac{n}{L}\right) = -L^{-1} \log\left(\frac{L}{n}\right).$$

Using Stirling's approximation we get  $f(\rho) = \rho \ln \rho + (1-\rho) \ln(1-\rho)$ . Using the fluctuation-dissipation relation we recover the well-known formula  $\sigma(\rho) = 2\rho(1-\rho)$  for the mobility in the case of SEP.

## D Derivation of the integral (130)

The result in (130) can be proved by comparing with the analysis of the time integrated current in a symmetric exclusion process on an infinite line [59, 73]. Consider an annealed initial state at uniform density  $\rho$ . Let  $Q_T$  be the time integrated current in a time window  $[0, T]$  through the site at origin. The current is related to the hydrodynamic density profile by

$$Q_T = \int_0^\infty dx \left[ \rho(x, T) - \rho(x, 0) \right].$$

On an average the current is zero because of the uniform initial density. However, all the even cumulants are non-zero. The analysis for the cumulant generating function of  $Q_T$  can be easily formulated in terms of the macroscopic fluctuation theory [59]. To analyze the variational formulation we take a series expansion method, similar to the one presented in Section 5. This way the fourth cumulant of current for the symmetric exclusion process can be related to that for non-interacting particles. The relation is similar to the one in (129):

$$\langle Q_T^4 \rangle_c = (1 - \alpha\rho)(1 - 2\alpha\rho)^2 \langle \hat{Q}_T^4 \rangle_c + 24\alpha\rho^2(1 - \alpha\rho)^2 \left[ -\mathcal{I} + \int_0^\infty dx \left( \frac{h(x, T)}{\rho^2} - \frac{h(x, 0)}{\rho^2} \right) \right], \quad (171)$$

where  $\mathcal{I}$  and  $\hat{h}$  are defined in section 5. We followed the same convention of denoting the non-interacting case by hat variables.

Using the Bethe ansatz Derrida and Gerschenfeld derived an exact expression for all the cumulants of the current for the SEP in the annealed setting [73]. The fourth cumulant in the annealed case can be extracted from their work:

$$\langle Q_T^4 \rangle_c = \rho(1 - \rho) \left[ 1 - 3\sqrt{2}\rho(1 - \rho) \right] \frac{2}{\sqrt{\pi}}. \quad (172)$$

On the other hand the cumulant for non-interacting particles corresponding to  $\alpha = 0$  can be derived by a direct counting argument [59] leading to

$$\langle \hat{Q}_T^4 \rangle_c = \rho \frac{2}{\sqrt{\pi}}. \quad (173)$$

Making use of these results we establish (130).

## E Function $G$

Here we prove an identity for the function  $G$  defined in equation (141). Let  $\gamma$  be a permutation of the set  $\Omega = \{y_{-n}, \dots, y_{-1}, y_1, \dots, y_n\}$  of  $2n$  elements. Note that the zeroth element has been excluded for convenience. The function  $G$  has the property that

$$\begin{aligned} \sum_{\gamma} G(-x, -y_{\gamma(-n)}, \dots, -y_{\gamma(-2)}, -y_{\gamma(-1)}) G(x, y_{\gamma(1)}, y_{\gamma(2)}, \dots, y_{\gamma(n)}) \\ = \sum_{\epsilon_{-n}} \dots \sum_{\epsilon_{-1}} \sum_{\epsilon_1} \dots \sum_{\epsilon_n} [\delta_{\sum_{j \neq 0} \epsilon_j, 0}] \prod_{\ell \neq 0} \frac{1}{2} \operatorname{erfc}\left(\epsilon_\ell \frac{x - y_\ell}{\sqrt{4t}}\right). \end{aligned} \quad (174)$$

The identity holds for any set of  $2n$  elements, with  $n$  being a positive integer, and the binary variables  $\epsilon_j = \pm 1$  for any  $j$ .

First, we want to prove

$$\sum_{\nu} G(x, y_{\nu(1)}, \dots, y_{\nu(n)}) = \prod_{j=1}^n \frac{1}{2} \operatorname{erfc}\left(\frac{x - y_j}{\sqrt{4t}}\right), \quad (175)$$

where  $\nu$  is a permutation of the set  $\{y_1, \dots, y_n\}$ . Using the definition of  $G$  in (141) we get

$$\sum_{\nu} G(x, y_{\nu(1)}, \dots, y_{\nu(n)}) = \sum_{\nu} \int_x^{\infty} dz_1 \int_{z_1}^{\infty} dz_2 \cdots \int_{z_{n-1}}^{\infty} dz_n \prod_{j=1}^n g_t(z_j | y_{\nu(j)}).$$

As  $z_j$ 's are dummy variables, the expression can be re-written as

$$\sum_{\nu} G(x, y_{\nu(1)}, \dots, y_{\nu(n)}) = \left[ \sum_{\nu} \int_x^{\infty} dz_{\nu(1)} \int_{z_{\nu(1)}}^{\infty} dz_{\nu(2)} \cdots \int_{z_{\nu(n-1)}}^{\infty} dz_{\nu(n)} \right] \prod_{j=1}^n g_t(z_j | y_j).$$

The range of integration denoted inside the square bracket is equal to the integration over the entire volume of all  $z_j \geq x$ , leading to

$$\sum_{\nu} G(x, y_{\nu(1)}, \dots, y_{\nu(n)}) = \prod_{j=1}^n \left[ \int_x^{\infty} dz_j g_t(z_j | y_j) \right] = \prod_{j=1}^n \frac{1}{2} \operatorname{erfc}\left(\frac{x - y_j}{\sqrt{4t}}\right), \quad (176)$$

proving (175).

In the next step we prove the identity (174) using (176). In (174) the summation is over all permutations of the set  $\Omega$ . Any ordered set which is generated by a permutation of  $\Omega$  can also be generated by suitably constructing two subsets of  $n$  elements each, and subsequently taking permutations within the elements of the subsets. To elaborate, let us divide the ordered set  $\Omega \equiv \{y_{-n}, \dots, y_{-1}, y_1, \dots, y_n\}$  into two ordered subsets  $\omega_L \equiv \{y_{-n}, \dots, y_{-1}\}$  and  $\omega_R \equiv \{y_1, \dots, y_n\}$ . The number of ordered sets generated by permutations of  $\Omega$  is  $(2n)!$ , all of which can also be generated by combination of the following steps.

1. The simplest are permutations strictly within  $\omega_L$  and  $\omega_R$ . There are  $n!n!$  such permutations.
2. One can construct another pair of ordered subsets  $\omega'_L$  and  $\omega'_R$  by exchanging any two elements between  $\omega_L$  and  $\omega_R$ . There are  $n^2$  pair of subsets that can be generated this way. Overall this produces  $n!n! \binom{n}{1}^2$  permutations.
3. We can further exchange two elements from  $\omega_L$  with two elements from  $\omega_R$ , and then consider permutations within those subsets. This generates  $n!n! \binom{n}{2}^2$  permutations.
4. Continuing this process we can generate all permutations  $\Omega$ .

As a useful check, one can compute all the permutations generated by the above procedure and find that it is equal to the total number of permutations:

$$n!n! \left[ 1 + \binom{n}{1}^2 + \binom{n}{2}^2 + \cdots + \binom{n}{n}^2 \right] = (2n)!$$

The above decomposition of permutations simplifies the summation on the left hand side in (174). For example, consider the ordered sets generated in step 1 of the decomposition where the ordered subsets are  $\nu_1 \omega_L$  and  $\nu_2 \omega_R$ , with  $\nu_1$  and  $\nu_2$  being permutation operators. Summing over all permutations we have

$$\sum_{\nu_1} G(-x, -y_{\nu_1(-n)}, \dots, -y_{\nu_1(-1)}) \sum_{\nu_2} G(x, y_{\nu_2(1)}, \dots, y_{\nu_2(n)}) = \prod_{j=1}^n \frac{1}{2} \operatorname{erfc}\left(-\frac{x - y_{-j}}{\sqrt{4t}}\right) \frac{1}{2} \operatorname{erfc}\left(\frac{x - y_j}{\sqrt{4t}}\right),$$

where we used the relation (175). We now introduce binary variables  $\{\epsilon_{-n}, \dots, \epsilon_{-1}, \epsilon_1, \dots, \epsilon_n\}$ , each of which take values  $\pm 1$ . Then the above formula can be re-written as

$$\sum_{\nu_1} G(-x, -y_{\nu_1(-n)}, \dots, -y_{\nu_1(-1)}) \sum_{\nu_2} G(x, y_{\nu_2(1)}, \dots, y_{\nu_2(n)}) = \prod_{j=-n, j \neq 0}^n \frac{1}{2} \operatorname{erfc}\left(\epsilon_j \frac{x - y_j}{\sqrt{4t}}\right),$$

with a configuration of the binary variables:  $\epsilon_j = -1$  for  $j < 0$  and  $\epsilon_j = 1$  for  $j > 0$ .

The summation over configurations generated in the step 2 is performed in a similar way. In this, the subsets are generated by exchange of one element each from  $\omega_L$  and  $\omega_R$ . For example, consider the  $(-n)$ th element from  $\omega_L$  and 1st element from  $\omega_R$  have been exchanged. The summation over configurations generated by permutation within these subsets yields a simple expression

$$\sum_{\nu_1} G(-x, -y_{\nu_1(1)}, -y_{\nu_1(-n+1)}, \dots, -y_{\nu_1(-1)}) \sum_{\nu_2} G(x, y_{\nu_2(-n)}, y_{\nu_2(2)}, \dots, y_{\nu_2(n)}) = \prod_{j=-n, j \neq 0}^n \frac{1}{2} \operatorname{erfc}\left(\epsilon_j \frac{x - y_j}{\sqrt{4Dt}}\right),$$

with  $\epsilon_{-n} = 1$ ,  $\epsilon_1 = -1$  and other binary variables are the same as before. From these examples the pattern emerges, namely the above decomposition of permutations leads to the identity (174).



## References

1. Chou T and Lohse D 1999 *Phys. Rev. Lett.* **82** 3552
2. Kärger J and Ruthven D 1992 *Diffusion in zeolites and other microporous solids* (New York: Wiley)
3. Das A, Jayanthi S, Deepak H S M V, Ramanathan K V, Kumar A, Dasgupta C and Sood A K 2010 *ACS Nano* **4** 1687
4. Gupta S, Rosso A and Texier C 2013 *Phys. Rev. Lett.* **111** 210601
5. Li G W, Berg O G and Elf J 2009 *Nature Physics* **5** 294
6. Hodgkin A L and Keynes R D 1955 *Journal of physiology* **128** 61
7. Richards P M 1977 *Phys. Rev. B* **16** 1393
8. De Masi A and Ferrari P A 1985 *J. Stat. Phys.* **38** 603
9. Kipnis C 1986 *Ann. Probab.* **14** 397
10. Majumdar S N and Barma M 1991 *Phys. Rev. B* **44** 5306
11. van Beijeren H 1991 *J. Stat. Phys.* **63** 47
12. Ferrari P A and Fontes L R G 1996 *J. Appl. Probab.* **33** 411
13. Liggett T 2004 *Interacting Particle Systems* (New York: Springer)
14. Gupta S, Majumdar S N, Godrèche C and Barma M 2007 *Phys. Rev. E* **76** 021112
15. Percus J 1974 *Phys. Rev. A* **9** 557
16. Roy A, Narayan O, Dhar A and Sabhapandit S 2013 *J. Stat. Phys.* **150** 851
17. Harris T E 1965 *J. Appl. Probab.* **2** 323
18. Arratia R 1983 *Ann. Probab.* **11** 362
19. Sethuraman S and Varadhan S R S 2013 *Annals of Probability* **41** 1461
20. Rödenbeck C, Kärger J and Hahn K 1998 *Phys. Rev. E* **57** 4382
21. Kollmann M 2003 *Phys. Rev. Lett.* **90** 180602
22. Imamura T and Sasamoto T 2007 *J. Stat. Phys.* **128** 799
23. Euán-Díaz E C, Misko V R, Peeters F M, Herrera-Velarde S and Castañeda Priego R 2012 *Phys. Rev. E* **86** 031123
24. Flomenbom O and Taloni A 2008 *EPL* **83** 20004
25. Lizana L and Ambjörnsson T 2008 *Phys. Rev. Lett.* **100** 200601
26. Ben-Naim E and Krapivsky P L 2009 *Phys. Rev. Lett.* **102** 190602
27. Manzi S J, Torrez Herrera J J and Pereyra V D 2012 *Phys. Rev. E* **86** 021129
28. Lizana L, Lomholt M A and Ambjörnsson T 2014 *Physica A* **395** 148
29. Krapivsky P L, Mallick K and Sadhu T 2014 *Phys. Rev. Lett.* **113** 078101
30. Barkai E and Silbey R 2009 *Phys. Rev. Lett.* **102** 050602
31. Kukla V, Kornatowski J, Demuth D, Girnus I, Pfeifer H, Rees L V C, Schunk S, Unger K K and Karger J 1996 *Science* **272** 702
32. Wei Q H, Bechinger C and Leiderer P 2000 *Science* **287** 625
33. Lutz C, Kollmann M and Bechinger C 2004 *Phys. Rev. Lett.* **93** 026001
34. Lin B, Meron M, Cui B, Rice S A and Diamant H 2005 *Phys. Rev. Lett.* **94** 216001
35. Siems U, Kreuter C, Erbe A, Schwierz N, Sengupta S, Leiderer P and Nielaba P 2012 *Scientific Reports* **2** 1015
36. Edwards S F and Wilkinson D R 1982 *Proc. Royal Soc. A* **381** 17
37. Kardar M, Parisi G and Zhang Y C 1986 *Phys. Rev. Lett.* **56** 889
38. Levitt D 1973 *Phys. Rev. A* **8** 3050
39. Alexander S and Pincus P 1978 *Phys. Rev. B* **18** 2011
40. Krapivsky P L, Redner S and Ben-Naim E 2010 *A Kinetic View of Statistical Physics* (Cambridge University Press)
41. Illien P, Bénichou O, Mejía-Monasterio C, Oshanin G and Voituriez R 2013 *Phys. Rev. Lett.* **111** 038102
42. Leibovich N and Barkai E 2013 *Phys. Rev. E* **88** 032107
43. Jara M and Landim C 2006 *Annales de l'Institut Henri Poincaré Probability and Statistics* **42** 567
44. Derrida B 2007 *J. Stat. Mech.* **P07023**
45. Touchette H 2009 *Phys. Rep.* **478** 1
46. Bertini L, De Sole A, Gabrielli D, Jona-Lasinio G and Landim C 2001 *Phys. Rev. Lett.* **87** 040601
47. Bertini L, De Sole A, Gabrielli D, Jona-Lasinio G and Landim C 2002 *J. Stat. Phys.* **107** 635
48. Bertini L, De Sole A, Gabrielli D, Jona-Lasinio G and Landim C 2005 *Phys. Rev. Lett.* **94** 030601
49. Bertini L, De Sole A, Gabrielli D, Jona-Lasinio G and Landim C 2009 *J. Stat. Phys.* **135** 857
50. Bertini L, De Sole A, Gabrielli D, Jona-Lasinio G and Landim C 2014 *ArXiv e-prints* **1404.6466**
51. Jordan A N, Sukhorukov E V and Pilgrim S 2004 *J. Math. Phys.* **45** 4386
52. Pilgram S, Jordan A N, Sukhorukov E V and Büttiker M 2003 *Phys. Rev. Lett.* **90** 206801
53. Derrida B 2011 *J. Stat. Mech.* **P01030**
54. Bertini L, Sole A D, Gabrielli D, Jona-Lasinio G and Landim C 2006 *J. Stat. Phys.* **123** 237
55. Bodineau T, Derrida B and Lebowitz J 2010 *J. Stat. Phys.* **140** 648
56. Bodineau T, Derrida B, Lecomte V and van Wijland F 2008 *J. Stat. Phys.* **133** 1013
57. Tailleur J, Kurchan J and Lecomte V 2007 *Phys. Rev. Lett.* **99** 150602
58. Bunin G, Kafri Y and Podolsky D 2012 *EPL* **99** 20002
59. Derrida B and Gerschenfeld A 2009 *J. Stat. Phys.* **137** 978
60. Jona-Lasinio G 2010 *Prog. Theor. Phys.* **124** 262
61. Jona-Lasinio G 2014 *J. Stat. Mech.* **P02004**
62. Krapivsky P L and Meerson B 2012 *Phys. Rev. E* **86** 031106
63. Meerson B and Sasorov P V 2013 *J. Stat. Mech.* **P12011**

- 
64. Vilenkin A, Meerson B and Sasorov P V 2013 *J. Stat. Mech.* **P06007**
  65. Meerson B and Sasorov P V 2014 *Phys. Rev. E* **89** 010101(R)
  66. Meerson B, Vilenkin A and Krapivsky P L 2014 *Phys. Rev. E* **90** 02210
  67. Hurtado P I, Espigares C P, del Pozo J J and Garrido P L 2014 *J. Stat. Phys.* **154** 214
  68. Krapivsky P L, Mallick K and Sadhu T 2015 *J Phys A: Math Theor* **48** 015005
  69. Krapivsky P L, Mallick K and Sadhu T 2015 *ArXiv e-prints* **1505.01287**
  70. Baek Y and Kafri Y 2015 *ArXiv e-prints* **1505.05796**
  71. Spohn H 1991 *Large Scale Dynamics of Interacting Particles* (New York: Springer-Verlag)
  72. Rajesh R and Majumdar S N 2001 *Phys. Rev. E* **64** 036103
  73. Derrida B and Gerschenfeld A 2009 *J. Stat. Phys.* **136** 1
  74. Freidlin M and Wentzell A 1984 *Random Perturbations of Dynamical Systems* (New York: Springer-Verlag)
  75. Martin P, Siggia E and Rose H 1973 *Phys. Rev. A* **8** 423
  76. De Dominicis C and Peliti L 1978 *Phys. Rev. B* **18** 353
  77. Hegde C, Sabhapandit S and Dhar A 2014 *Phys. Rev. Lett.* **113** 120601
  78. Elgart V and Kamenev A 2004 *Phys. Rev. E* **70** 041106
  79. Prudnikov A P, Brychkov Yu A and Marichev O I 1986 *Integrals and Series: Special Functions* vol 2 (New York: Gordon & Breach Science Publishers)
  80. Kumar D 2008 *Phys. Rev. E* **78** 021133
  81. Burlatsky S, Oshanin G, Mogutov A and Moreau M 1992 *Phys. Lett. A* **166** 230
  82. Landim C, Olla S and Volchan S B 1998 *Commun. Math. Phys.* **192** 287
  83. Gärtner J and Presutti E 1990 *Ann. Inst. Henri Poinc.* **53** 1
  84. Ferrari P A 1991 *Ann. Inst. Henri Poinc.* **55** 637
  85. Alexander F J, Janowsky S A, Leibowitz J L and van Beijeren H 1993 *Phys. Rev. E* **47** 403
  86. Babelon O, Bernard D and Talon M 2003 *Introduction to Classical Integrable Systems* (Cambridge: Cambridge University Press)
  87. Gradshteyn I S and Ryzhik I M 2007 *Table of Integrals, Series, and Products* (London: Academic Press)